# INVESTIGACION y CIENCIA

Edición española de SCIENTIFIC AMERICAN

EL PROBLEMA INFORMATICO DEL AÑO 2000

INACTIVACION DE LOS VIRUS DE LA GRIPE

INFORME ESPECIAL: LA COSMOLOGIA A DEBATE




MARZO 1999
800 PTA. 4,81 EURO

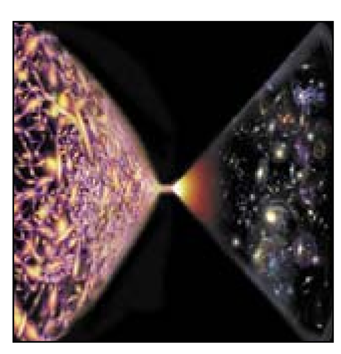



Marzo de 1999 Número 270

Edición española de SCIENTIFIC AMERICAN



## SECCIONES



**Portada:** Don Dixon

## PROCEDENCIA DE LAS ILUSTRACIONES

| Página | Fuente |
|---|---|
| 8-10 | Peter Challis |
| 11 | Daniels & Daniels |
| 12 | Daniels & Daniels (*izquierda*); David Schneider (*gráficas*) |
| 13 | David Schneider, fuente: Craig J. Hogan, Robert P. Kirshner y Nicholas B. Suntzeff |
| 14 | P. Ruiz-Lapuente |
| 15-16 | Proyecto Supernovas y Cosmología |
| 17 | Observatorio del Roque de los Muchachos (*izquierda*), P. Ruiz-Lapuente, J. Mendez y N. Walton (*derecha*) |
| 18 | P. Ruiz-Lapuente |
| 22-23 | Alfred T. Kamajian |
| 24 | Archivos de Einstein, Universidad de Zurich |
| 25 | Steve K. Lamoreaux (*fotografía*), Jared Schneidman Design |
| 27 | Don Dixon (*ilustración*), Dmitry Krasny (*dibujo*) |
| 28 | George Musser y Dmitry Krasny |
| 30-31 | Alfred T. Kamajian |
| 32-33 | Dmitry Krasny |
| 34 | Don Dixon, fuente: Martin A. Bucher y David N. Spergel |
| 35 | Stuart Levy y Tamara Munzner |
| 37 | Don Dixon |
| 52-53 | Ken Preston-Mafham |
| 54 | Ken Preston-Mafham, Douglas W. Tallamy (*arriba, a la izquierda y centro*) |
| 55-56 | Ken Preston-Mafham |
| 57 | C. Allan Morgan |
| 58-59 | Bryan Christie |
| 60-61 | Thomas Brock, cortesía de Michael Madigan (*fotografía*), Bryan Christie |
| 62 | Stuart Butterworth y W. Graeme Laver |
| 63 | Bryan Christie, fuente: R. Scott Rowland |
| 64 | David Hockley (*izquierda*), Richard Compans (*derecha*) |
| 65 | Stephen Rose |
| 66-67 | Corbis-Bettmann (*izquierda*), Robyn Beck (*derecha*) |
| 68-69 | Tomo Narashima |
| 70 | Tomo Narashima y George Retseck |
| 71 | Nigel J. Dennis |
| 72-73 | Biophoto Associates/Science Source (*arriba, izquierda*), Parviz M. Pour (*arriba, derecha*), George Retseck (*ilustraciones*) |
| 75-76 | Slim Films, Scott Camazine (*colmena*), Scott Smith y Sandro Vannini (*abejas*) |
| 77 | Chip East |
| 78 | Bryan Christie |
| 79 | Chip East |
| 80 | Bryan Christie |
| 84 | Matt Collins |


**COLABORADORES DE ESTE NUMERO**

**Asesoramiento y traducción:**

Pilar Ruiz-Lapuente: *Exploración del espacio-tiempo mediante supernovas*; Juan P. Campos: *Antigravedad cosmológica*; Jaume Garriga: *Inflación en un universo de baja densidad*; Luisa Vilaplana: *Cuidado de la prole entre los insectos*; José M.ª Valderas Martínez: *Desarme de los virus de la gripe* y *Nexos*; Esteban Santiago: *Microsatélites de ADN*; Luis Bou: *El problema del año 2000* y *Juegos matemáticos*; Angel Garcimartín: *Perfiles*; J. Vilardell: *Hace..., Taller y laboratorio* e *Ideas aplicadas*.




**SUSCRIPCIONES**

Prensa Científica S. A.
Muntaner, 339 pral. 1.ª
08021 Barcelona (España)
Teléfono 93 414 33 44
Fax 93 414 54 13

**Precios de suscripción:**

| | Un año | Dos años |
|---|---|---|
| España | 8.800 pta. | 16.000 pta. |
| | 52,89 euro | 96,16 euro |
| Extranjero | 11.150 pta. | 20.700 pta. |
| | 67,01 euro | 124,41 euro |

**Ejemplares sueltos:**

Ordinario: 800 pta. 4,81 euro
Extraordinario: 1.000 pta. 6,01 euro

—El precio de los ejemplares atrasados es el mismo que el de los actuales.

**DISTRIBUCION**

**para España:**

**MIDESA**
Aragoneses, 18 (Pol. Ind. Alcobendas)
28108 Alcobendas (Madrid)
Tel. 91 484 39 00

**para los restantes países:**

Prensa Científica, S. A.
Muntaner, 339 pral. 1.ª – 08021 Barcelona
Teléfono 93 414 33 44

**PUBLICIDAD**

GM Publicidad
Francisca Martínez Soriano
Menorca, 8, semisótano, centro, izquierda.
28009 Madrid
Tel. 91 409 70 45 – Fax 91 409 70 46

Cataluña y Baleares:
Miguel Munill
Muntaner, 339 pral. 1.ª
08021 Barcelona
Tel. 93 321 21 14
Fax 93 414 54 13

Difusión controlada





ISSN 0210136X Dep. legal: B. 38.999 – 76

Filmación y fotocromos reproducidos por Dos Digital, Zamora, 46-48, 6ª planta, 3ª puerta - 08005 Barcelona
Imprime Rotocayfo, S.A. Ctra. de Caldes, km 3 - Santa Perpètua de Mogoda (Barcelona)

Printed in Spain - Impreso en España

# HACE...

## ...cincuenta años

**EL INFLUYENTE EINSTEIN.** «Albert Einstein, cuyo 70 cumpleaños se señala para este mes en todo el mundo civilizado, ocupa un lugar singular entre los científicos. Débese su fama a la relatividad, sin duda. El retrato que ofrecemos data del año de su mayor productividad, 1905. Mientras trabajaba de empleado en la oficina de patentes de Suiza, realizó su gran contribución a la teoría cuántica y desarrolló la teoría de la relatividad restringida.»

**REACCIÓN A LA TENSIÓN NERVIOSA.** «La experimentación sobre el 'sistema de adaptación general' ha llevado al doctor Hans Seyle, de la Universidad de Montreal, a formular la siguiente hipótesis ya admitida: Una tensión psíquica de larga duración provoca una producción excesiva de hormona estimuladora de la adrenalina en la pituitaria anterior; ello obliga a la corteza suprarrenal a una descarga intensiva de hormonas parecidas a la desoxicorticosterona, lo cual, entre otras cosas, afecta al riñón haciendo que éste libere sustancias hipertensoras. Si investigaciones ulteriores demuestran que la tensión nerviosa crónica puede producir los mismos trastornos en el hombre que en los animales, cabría inferir que las enfermedades más frecuentes y letales de hoy se deben al 'cansancio y prisas' de la vida moderna.»

## ...cien años

**POLÍTICA Y ASTRONOMÍA.** «Los grandes observatorios del mundo se encuentran en las proximidades de grandes ciudades o universidades importantes (lugares elegidos por motivos locales o políticos) donde no resultan aptos para investigaciones astronómicas rigurosas. Apartarse de tal precedente suponía dar un paso valiente. Ese salto lo ha dado una mujer, miss Catherine Bruce, de Nueva York, quien donó 50.000 dólares al observatorio de la facultad de Harvard. El telescopio fotográfico Bruce está instalado en Arequipa (Perú), en una región inmejorable para la observación. Con su ayuda, se han descubierto nuevas estrellas en la Gran Nube de Magallanes que muestran una conexión adicional de ese objeto con la Vía Láctea.»

**SUMINISTRO VITAL.** «Monsieur Georges Jaubert ha estado experimentando con el abastecimiento de aire para el consumo de un hombre encerrado en un recinto hermético, como una campana de buceo. Planteó que el 79 por ciento del nitrógeno contenido en el aire respirable permanece intacto tras haberse consumido el 21 por ciento del oxígeno, y que ese mismo nitrógeno mezclado con un suministro de oxígeno fresco se hace respirable cuando se eliminan el dióxido de carbono y el vapor de agua producidos por el aliento. Halló que tal hipótesis es correcta; asimismo descubrió una sustancia química que en contacto con la atmósfera limpia todo el aire viciado de los gases impuros generados por la respiración.»

**AMIGOS, ROMANOS.** «En las nuevas excavaciones del Foro Romano, un hallazgo de inigualable interés es la basa de la columna erigida en el punto donde fue incinerado el cadáver de César. Suetonio habla de una columna de mármol de Numidia dedicada *parenti patriae* en ese lugar. También se levantó allí un altar, pero fue destruido ya que el culto a César era ilegal. Posteriormente, Augusto construyó allí el Templo de Julio. Delante del podio del templo hay un nicho semicircular, donde, sobre un pavimento de adoquines travertinos, quedan los restos de una basa como la que cabría esperar que tuviera la columna. Allí descansó su cuerpo. Allí Antonio excitó a la plebe; y allí, de las cenizas de fénix de una república muerta se alzó el joven imperio.»



*Albert Einstein en 1905*

## ...ciento cincuenta años

**INVENCIÓN DE LA INCURSIÓN AÉREA.** «'La Prensa', de Viena (Austria), informa de lo siguiente: 'Venecia va a ser bombardeada desde globos, ya que los canales impiden el acceso a la artillería. En Treviso se están preparando cinco globos, de siete metros de diámetro cada uno. Con un viento favorable serán lanzados y dirigidos lo más cerca posible de Venecia, y al ser llevados a una posición vertical sobre la ciudad, serán disparados por electromagnetismo merced a un largo cable de cobre aislado unido a una gran batería galvánica instalada en la orilla. Cada bomba cae perpendicularmente y estalla al llegar a tierra.'» [*Nota de la Redacción: Fue este el primer bombardeo aéreo, y parece que tuvo éxito, ya que la revuelta veneciana se deshizo.*]

**LUJO.** «Lyon es el centro de la gran región francesa manufacturera de seda. Su población cercana a las 200.000 personas se agita en casas altas y desiguales que se hacinan y ensombrecen unas calles estrechas, tortuosas y mugrientas. No hay grandes talleres similares a nuestros telares algodoneros. Todo se hace bajo el techo familiar. La vida es mísera; la población dedicada a la tejeduría vive en una depravación espantosa. Por unos pocos céntimos diarios, esos desdichados, cansados y hambrientos, trabajan duramente para el adorno de quienes no creen que en el mundo exista la pobreza extrema.»



INFORME ESPECIAL

# REVOLUCION EN LA COSMOLOGIA

En un congreso celebrado en mayo de 1998 sobre "La energía perdida del universo" los cosmólogos tuvieron que tomar partido. ¿Creían en esas recientes observaciones de estrellas que habían estallado, según las cuales —y contra todo pronóstico— el universo crece con celeridad creciente? Se sabe desde los años veinte que el tamaño del universo aumenta y las galaxias se separan más y más, pero siempre se había supuesto que la expansión se moderaba, refrenada por la fuerza atractiva de la gravedad. Si el crecimiento se acelera, es que el universo está lleno de alguna forma de materia o de energía desconocida cuya gravedad, en vez de atraer, repele. Pero la idea de una forma de energía inédita repugna a los físicos. Y, sin embargo, de los 60 allí presentes, 40 aceptaron los nuevos hallazgos.

Hacía ya más de diez años que se venía sospechando que las cosas no acababan de encajar en la cosmología moderna. Cuando se echan las cuentas de la materia ordinaria que hay en el universo no nos sale la que se precisa para frenar la expansión cósmica, según lo predicho por la teoría de la inflación, modelo muy vertebrado de las primeras etapas de la gran explosión. Hasta ahora, las pruebas aducidas contra esa teoría no ponían en peligro sus ventajas reconocidas. Las cosas han dado un vuelco. Algo falla. Como muy poco, la expansión no se decelera al ritmo esperado. Si no transigen con una energía extravagante, los astrofísicos habrán de abandonar o modificar la inflación.

INVESTIGACIÓN Y CIENCIA presenta en este número tres caras del problema. En los dos primeros artículos, los dos programas de investigación sobre supernovas exponen los resultados que han causado semejante conmoción. A continuación, un teórico explica por qué esos resultados hablan en favor de que el espacio vacío está lleno de una energía etérea. Por último, un par de cosmólogos ofrece otra interpretación, que extiende la teoría de la inflación a tiempos "anteriores" a la gran explosión.

*—La redacción*

# Exploración del espacio-tiempo

*Las explosiones de estrellas observadas a inmensas distancias muestran que la expansión cósmica podría estar acelerándose, lo que nos lleva a pensar en la intervención de una nueva forma de energía, exótica*

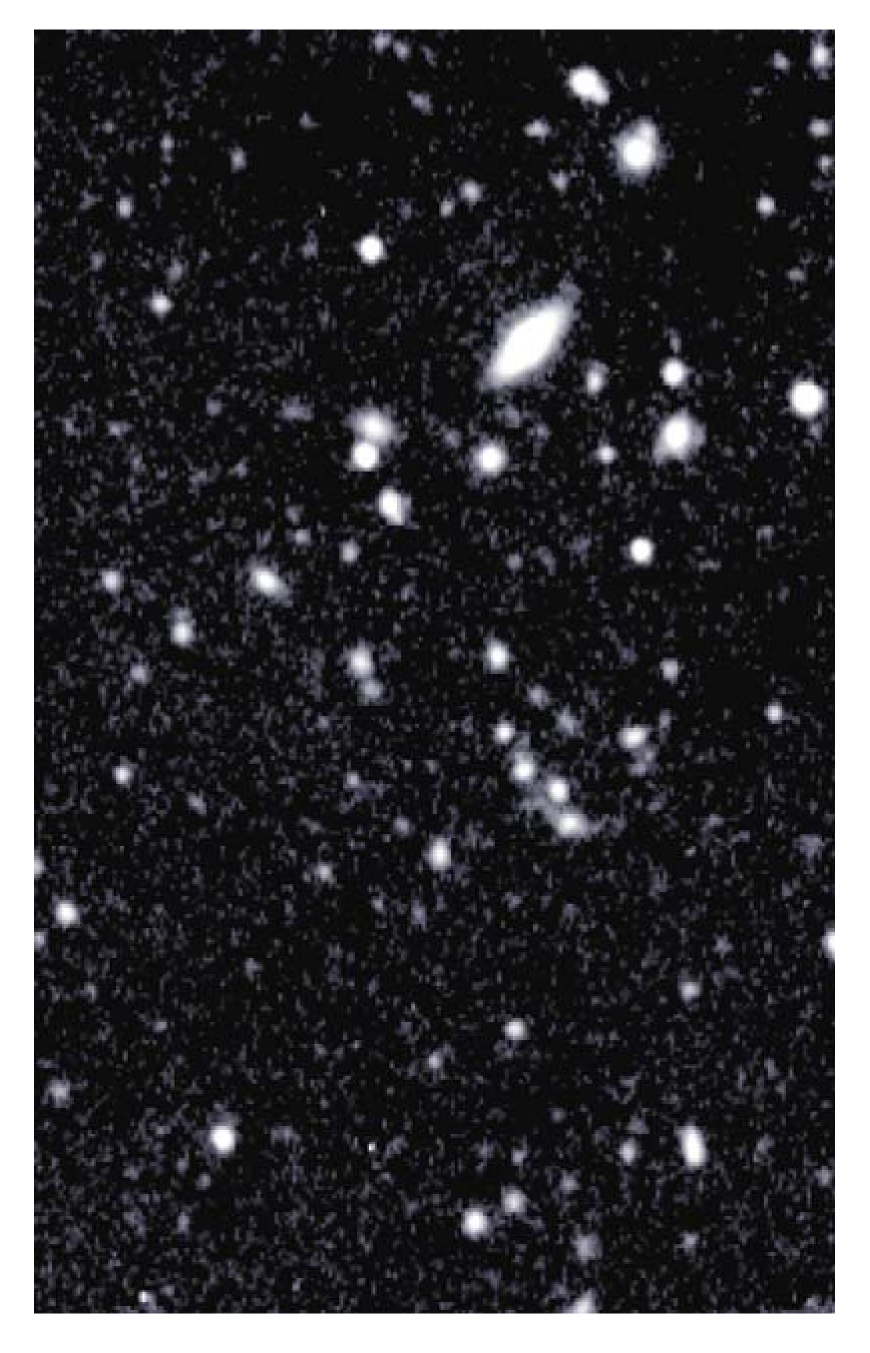
**Craig J. Hogan, Robert P. Kirshner y Nicholas B. Suntzeff**

Hace mucho tiempo (unos 5000 millones de años), en una galaxia muy lejana (a unos 2000 megaparsecs) explotó una estrella y su destello brilló más que miles de millones de soles juntos. Su luz se propagó por el espacio, apagándose y aumentando en longitud de onda a través de un cosmos en expansión, antes de que una brizna de ella llegara a la Tierra. En un intervalo de 10 minutos de cierta noche oscura de 1997, unos centenares de fotones de la supernova, no más, aterrizaron en el espejo de un telescopio de Chile. En el observatorio, un ordenador creó una imagen digital donde se evidenciaba el puntito luminoso. Pese a su apariencia irrelevante, aquella débil mota fue para nosotros una ráfaga de luz, un faro que alumbraba nuestra exploración del espacio y el tiempo.

En colaboración con otros compañeros repartidos por diversos países, hemos seguido la llegada de la luz emitida desde docenas de supernovas. Tales observaciones nos han facultado para cartografiar la forma del universo y establecer un esbozo cronológico de su expansión. La conclusión a la que nosotros y otro equipo de astrónomos hemos arribado pone en tela de juicio postulados aceptados desde hace decenios. Según parece, el universo es mayor y está más vacío de lo esperado. Además, su expansión no remite, como muchos cosmólogos habían pronosticado, sino que se acelera.

### Curvas estelares

A lo largo de buena parte del siglo se ha mantenido vivo el interés en torno al curso de la expansión cósmica. Por un doble motivo: da cuenta de la geometría del universo y refleja la naturaleza de sus constituyentes, es decir, de la materia, la luz y, posiblemente, otras formas de energía más sutiles. La teoría de la relatividad general de Albert Einstein aúna estas propiedades fundamentales del universo y describe la manera en que condicionan el movimiento de la materia y la propagación de la luz; merced a ello, puede predecir determinados fenómenos que los astrónomos someten a medición.



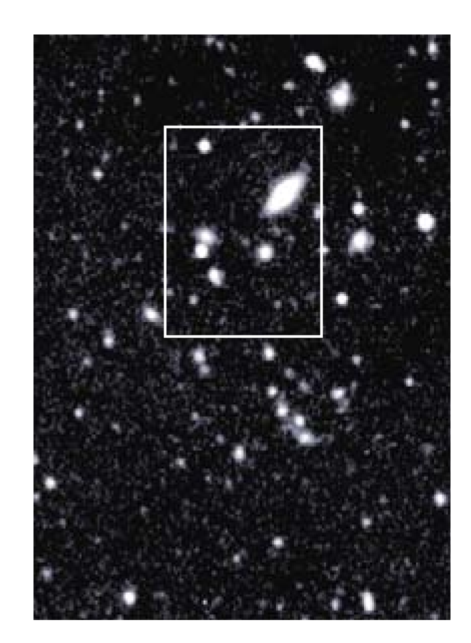
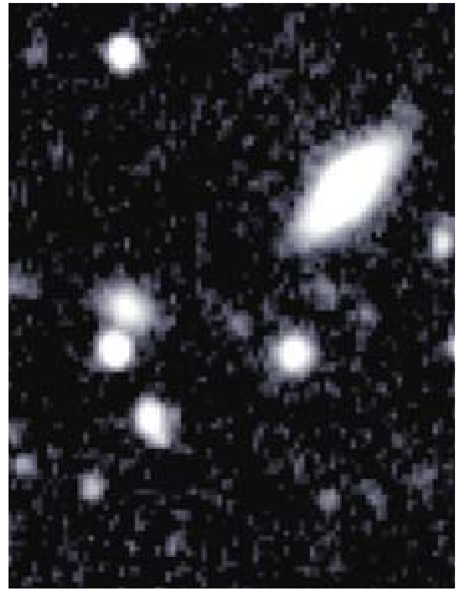
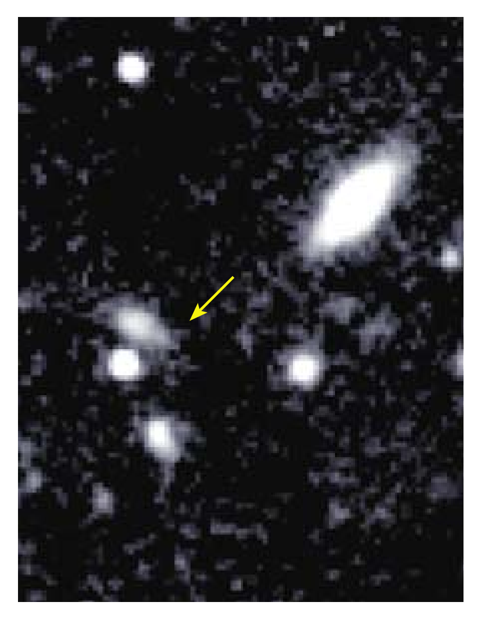
# mediante supernovas

Antes de que Einstein publicara su teoría en 1916 y de las primeras observaciones de la expansión cósmica, acontecidas durante la década siguiente, se creía que el universo mantenía inalterado su tamaño. El propio Einstein desconfió de sus ecuaciones cuando advirtió que implicaban un universo dinámico. La situación cambió, sin dejar lugar a la duda, con las nuevas mediciones de los movimientos de las galaxias realizadas por Edwin P. Hubble y otros: las galaxias lejanas y débiles se alejaban de la Tierra más deprisa que las galaxias cercanas y brillantes, tal como predecía la relatividad general para un universo que se expande y distancia, por ende, unas galaxias de otras. Los astrónomos en cuestión midieron la velocidad de alejamiento de las galaxias a través del corrimiento de líneas espectrales visibles hacia longitudes de onda mayores (el corrimiento al rojo). Aunque suele asociársele con el efecto Doppler —fenómeno responsable del cambio de tono del silbato de un tren que pasa o de una bocina de coche—, el corrimiento al rojo cosmológico debe entenderse, con mayor rigor, como el resultado de la expansión del universo; ésta alarga la longitud de onda de la luz que viaja entre las galaxias. Las emisiones procedentes de objetos remotos, al haber viajado durante más tiempo, experimentan un corrimiento al rojo mayor que la radiación de fuentes próximas.

El estado de la técnica en tiempos de Hubble limitó la comprobación de la expansión cósmica a galaxias que estaban bastante cerca. En el intervalo temporal que invirtió la luz en llegar a la Tierra desde esas galaxias próximas, el universo sólo se había expandido una pequeña fracción de su tamaño global. Con estos cambios modestos, el corrimiento al rojo es directamente proporcional a la distancia; el cociente fijo de los dos se llama constante de Hubble y describe el ritmo actual de expansión cósmica. Desde hace años, los astrónomos esperaban comprobar que las galaxias más alejadas se apartan de esta relación simple entre corrimiento al rojo y distancia, ya sea porque el ritmo de expansión ha cambiado a lo largo del tiempo o porque el espacio que media esté curvado. Medir ese efecto constituye, pues, un objetivo crucial para los cosmólogos; es un reto difícil que exige medios para acotar las inmensas distancias a que se encuentran las galaxias.

Hubble y otros calcularon las distancias a diversas galaxias partiendo del supuesto de que todas ellas tenían el mismo brillo intrínseco. De acuerdo con su razonamiento, las que lucían mucho se hallaban cerca y, lejos, las de brillo pálido. Pero ese planteamiento resulta válido sólo a grandes rasgos, pues las galaxias difieren en sus propiedades. Y falla por completo cuando hablamos de fuentes lejanas cuya luz tarda mu-

**1. ¿DONDE ESTA LA SUPERNOVA? Este par de imágenes, realizadas con el Telescopio Blanco de cuatro metros de diámetro del Observatorio Interamericano de Cerro Tololo, aportaron la primera prueba de una supernova. En la imagen de la derecha, obtenida tres semanas más tarde que la imagen de la izquierda, la supernova altera sutilmente la apariencia de una de las galaxias, aunque se percibe. ¿Puede el lector encontrarla? Algunas discrepancias se deben a las condiciones atmosféricas cambiantes. Solución: en la figura siguiente.**

**2. UNA SUPERNOVA REMOTA, con un corrimiento al rojo de $z = 0{,}66$, aparece junto a la flecha. La explosión de esta estrella afecta sólo a algunos elementos de la imagen obtenida después del suceso.**

chísimo en alcanzar la Tierra y nos muestra el estado de esas galaxias distantes hace miles de millones de años (de jóvenes); el brillo intrínseco de esas galaxias remotas pudo haber diferido notablemente del que emiten galaxias más viejas que observamos en nuestra vecindad. Es difícil desentrañar de los efectos de la expansión tales cambios evolutivos. Por eso, los astrónomos han venido buscando otras “candelas estándar” cuyo brillo intrínseco sea mejor conocido.

Para verse hoy lo que lució hace miles de millones de años luz, los faros emisores deben ser muy potentes. Al comienzo de los setenta, algunos se fijaron en los cuásares, fuentes portadoras de inmensa energía (probablemente alimentadas por agujeros negros que atrapan a las estrellas y gas de su alrededor). Pero los cuásares estudiados divergían unos de otros incluso más que las galaxias; resultaron, pues, de escasa utilidad.

Pero otros pensaron en las supernovas. Esas estrellas que explotan podían, quizá, servir de candelas estándar para la exploración cosmológica. Se trataba de un camino controvertido, pues las supernovas manifestaban también una amplia variación en sus propiedades. Sin embargo, la investigación realizada, a lo largo de los últimos diez años, por nuestro equipo ha permitido determinar, con exquisita precisión, el brillo intrínseco de las supernovas de tipo Ia.

## Muerte estelar

¿En qué consiste una supernova de tipo Ia? A grandes trazos, es el estallido final producido cuando una estrella que ha muerto se transforma en una bomba termonuclear natural. Pese a la espectacularidad de la transformación postrera, el comienzo de la vida del progenitor es el de una estrella ordinaria, una bola estable de gas cuyas capas externas se mantienen por el calor de las incesantes reacciones nucleares que se desarrollan en su interior y convierten hidrógeno en helio, carbono, oxígeno, neón y otros elementos. Cuando una estrella muere, las cenizas nucleares se mezclan en ascuas incandescentes, comprimidas por la gravedad hasta alcanzar el tamaño de la Tierra y la densidad de un millón de veces la típica de la materia ordinaria.

Esas estrellas enanas blancas, en su mayoría, se enfrían y apagan, extinguiéndose. Ahora bien, si ocurre que se halla en órbita cerca de otra estrella, puede succionar materia de su compañera y adquirir una densidad creciente, hasta que prende una ignición termonuclear descontrolada. El cataclismo nuclear destroza por entero la estrella enana, lanzando material a unos 10.000 kilómetros por segundo. Y el brillo de esta bola de fuego en expansión tarda unas tres semanas en alcanzar su máximo, para declinar en meses.

Aunque varían ligeramente en el brillo, estas supernovas siguen un patrón, en el que las explosiones mayores y más brillantes duran más que las débiles. Por tanto, si investigamos su duración, podremos corregir las diferencias y deducir su brillo intrínseco con una precisión de hasta el 12 %. En el curso de los últimos diez años, el estudio de supernovas de tipo Ia con los detectores modernos ha convertido a tales destellos de luz en las candelas estándar mejor conocidas de los astrónomos.

La frecuencia con que una luz de ésas aparece en una galaxia típica se cifra en torno a una vez cada 300 años. En la Vía Láctea son sucesos poco habituales. Si escudriñamos algunos miles de galaxias descubriremos, casi cada mes, una supernova de tipo Ia. Hay tantas galaxias en el universo que, cada pocos segundos, estallan en el firmamento supernovas cuyo brillo nos las hace accesibles al estudio. Sólo queda descubrirlas e investigarlas. A esa tarea se ha aplicado nuestro grupo, “High-Z Team” (donde Z designa el símbolo de corrimiento al rojo); el equipo se organizó en 1995 bajo la dirección de Brian P. Schmidt, de los observatorios de Mount Stromlo y Siding Spring en Australia. Compite en este campo el grupo formado en torno al programa Supernova, “Supernova Cosmology Project”, que empezó en 1988 y está liderado por Saul Perlmutter, del Laboratorio Nacional Lawrence en Berkeley (véase *Supernovas y expansión acelerada del universo*, de Pilar Ruiz Lapuente, Alex G. Kim y Nicholas Walton en este mismo número).

Cada equipo desarrolla su propio programa. Ambos, sin embargo, sacan partido de un mismo avance fundamental, el despliegue de grandes detectores electrónicos de luz en telescopios gigantes. Esta combinación produce imágenes digitales de objetos débiles sobre una superficie considerable de cielo. Los dos grupos han compartido, por dar un ejemplo de los nuevos medios técnicos, la Cámara de Gran Rendimiento, desarrollada por Gary M. Bernstein y J. Anthony Tyson. Cuando esta cámara se coloca en el foco del Telescopio Blanco, de cuatro metros, del Observatorio Interamericano de Cerro Tololo, con una sola exposición se cubre casi tanta extensión como la luna llena y obtiene una imagen de unas 5000 galaxias en 10 minutos.

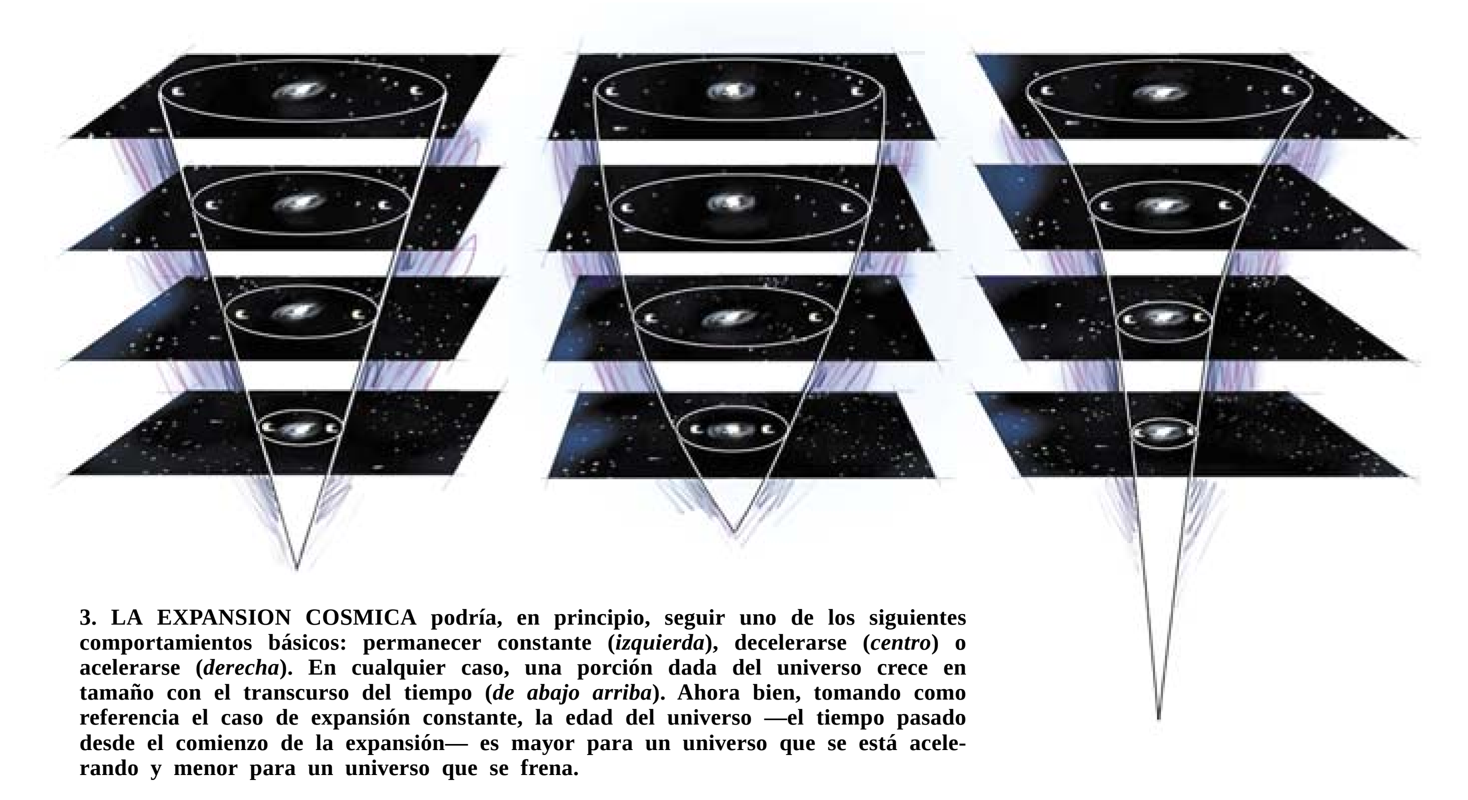
Para encontrar supernovas distantes, no hay más que tomar imágenes de la misma fracción del firmamento con unas pocas semanas de diferencia y buscar, en los cambios operados, explosiones de estrellas. Los detectores digitales cuentan el número de fotones en cada elemento de imagen de forma precisa. Por tanto, a nosotros sólo nos queda sustraer la primera imagen de la segunda y comprobar las diferencias significativas entre las dos. Puesto que se examinan miles de galaxias en cada par de imágenes, podemos estar seguros de que la búsqueda entre numerosas parejas de imágenes terminará por hallar muchas supernovas, siempre que el tiempo de observación acompañe. Para nuestra fortuna, la situación del observatorio, al pie de los Andes en la cordillera sur del desierto chileno de Atacama (uno de los lugares más secos del mundo), asegura cielos despejados. Con la esperanza de cosechar resultados, preparamos con antelación el tiempo de observación que debía desarrollarse en una batería de telescopios dispersos por todo el mundo, para que así las mediciones de seguimiento se acometieran antes de que se apagaran las supernovas.

La búsqueda de estrellas que explotan en los cielos desata un terremoto de actividad frenética. Los astrónomos han de tomar y comparar cientos de imágenes digitales a un ritmo vertiginoso. En todas las estancias de Cerro Tololo hay ordenadores que ejecutan la orden impuesta de alinear imágenes, corregir diferencias en transparencia atmosférica y tamaño de la imagen y sustraer los dos registros. En condiciones de normalidad, la mayoría de las galaxias desaparecen, dejando sólo un "ruido" en la diferencia entre dos imágenes. Las señales mayores denuncian un objeto nuevo o cambiante que puede ser una estrella variable, cuásares, asteroides y, en contados casos, supernovas.

Nuestros programas informáticos registran la posición de los nuevos objetos y se aprestan a identificar supernovas. Pero los tests automatizados son imperfectos. Hemos de escrutar las imágenes de forma visual para determinar si una supuesta supernova lo es en verdad. Puesto que de nuestros descubrimientos se ha de realizar un seguimiento inmediato en otros telescopios, se pasa al análisis sin perder tiempo.

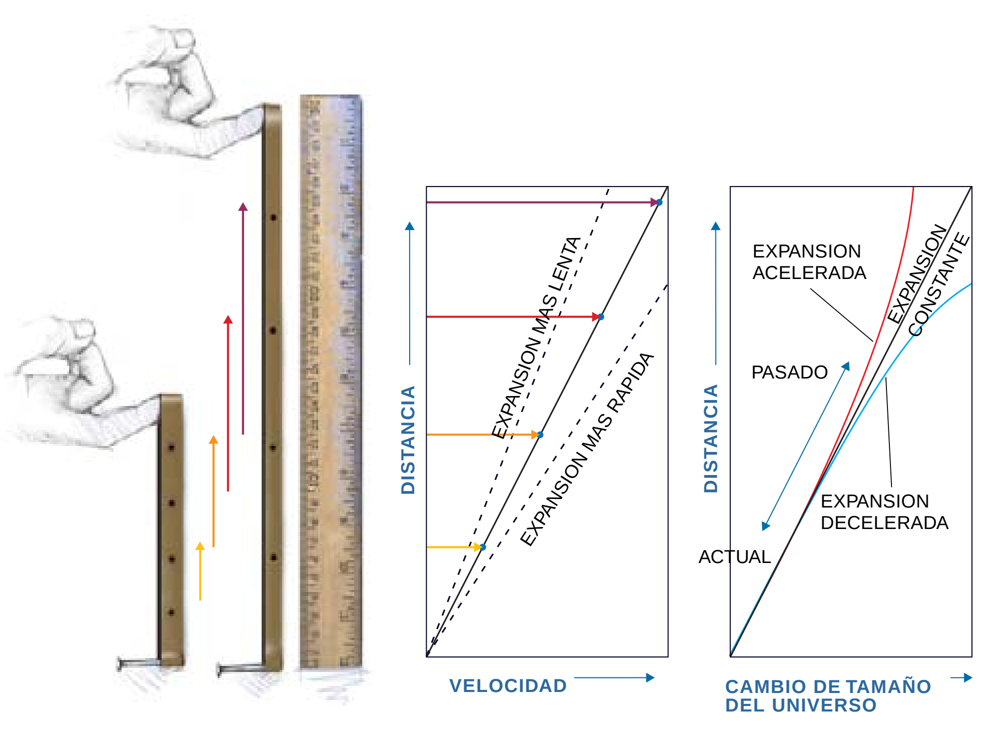
A continuación, nos centramos en los mejores candidatos de supernova. Los observamos con los mayores instrumentos ópticos del mundo, los telescopios Keck recién construidos en Hawai. Tales observaciones habrán de resolver si los objetos descubiertos son o no supernovas de tipo Ia, calibrar con exactitud su brillo intrínseco y determinar su corrimiento al rojo.

CRAIG J. HOGAN, ROBERT P. KIRSHNER y NICHOLAS B. SUNTZEFF comparten un interés por las grandes explosiones. Hogan, doctor por la Universidad de Cambridge, dirige el departamento de astronomía de la de Washington. Kirshner se doctoró en el Instituto de Tecnología de California, estudiando una supernova de tipo Ia observada en 1972 (la más brillante desde 1937). Codirige el Centro Smithsoniano de Astrofísica en Harvard. Suntzeff, formado en la Universidad de California en Santa Cruz, trabaja en el Observatorio Interamericano de Cerro Tololo.



**3. LA EXPANSION COSMICA podría, en principio, seguir uno de los siguientes comportamientos básicos: permanecer constante (*izquierda*), decelerarse (*centro*) o acelerarse (*derecha*). En cualquier caso, una porción dada del universo crece en tamaño con el transcurso del tiempo (*de abajo arriba*). Ahora bien, tomando como referencia el caso de expansión constante, la edad del universo —el tiempo pasado desde el comienzo de la expansión— es mayor para un universo que se está acelerando y menor para un universo que se frena.**

**4. EL EXPERIMENTO de la cinta elástica muestra la relación lineal entre velocidad de recesión y lejanía. Se ofrecen dos instantes de la cinta estirada hacia arriba a cierto ritmo. La velocidad de los puntos marcados en la goma está dada por la longitud de las flechas de colores. El punto más cercano al origen es el que menos se mueve durante el intervalo entre los dos instantes; por tanto, su velocidad es la menor (*flecha amarilla*). El punto más alejado es el que más se mueve; por tanto, su velocidad es la más alta (*flecha violeta*). La pendiente de la línea resultante es el ritmo de expansión (*gráfico de la izquierda*). Si el ritmo cambia con el tiempo, la pendiente lo hará también (*gráfico de la derecha*). Los tiempos iniciales aparecen hacia arriba a la derecha, porque la luz de los objetos alejados tarda más tiempo en alcanzar la Tierra, el origen del dibujo. Si el ritmo era más lento en el pasado —señal de que la expansión se acelera— la línea se curvará hacia arriba (*línea roja*). Si el ritmo era más rápido —en un universo que decelera su expansión— la línea continuará curvándose hacia abajo (*línea azul*).**




## En el lado oscuro

Mientras tanto, otros componentes del grupo, que tra- bajan con telescopios de Australia, Chile y los EE.UU., siguen la evolución de las supernovas para determinar el momento en que alcanzan el máximo de brillo y su lento desvanecimiento. Dura meses la campaña de observación de una supernova. A menudo, el análisis final tiene que esperar un año o más, intervalo en que ha desaparecido la luz de la estrella que ha estallado y podemos obtener una buena imagen de su galaxia huésped. Nos sirve esta vista final para sustraer, de las imágenes de la supernova, el brillo constante de la galaxia. Las mejores mediciones que hemos logrado proceden del *Telescopio Espacial Hubble*, capacitado para registrar detalles sutilísimos que nos permiten separar la estrella que ha explotado de su galaxia huésped.

Los dos equipos han estudiado ya ejemplos de supernova en alto corrimiento al rojo, que estallaron hace entre 4000 y 7000 millones de años. Tenía entonces el universo entre la mitad y dos tercios de su edad actual. En el curso de tales estudios, uno y otro grupo recibieron una misma sorpresa, mayúscula: las supernovas son más débiles de lo esperado. La discrepancia entre lo esperado y lo observado es leve; en promedio, las supernovas remotas aparecen un 25 por ciento más débiles de lo esperado. Ahora bien, basta esa cifra para cuestionar las teorías cosmológicas más arraigadas.

Antes de extraer ninguna conclusión radical, los astrónomos de ambos equipos han buscado explicaciones más llanas que justificaran la relativa debilidad de las supernovas remotas. Se apela a la oscuridad provocada por el polvo cósmico, que apantallaría parte de la luz. Pero nosotros descartamos esa posibilidad por la sencilla razón de que los granos de polvo filtrarían la luz azul más que la roja, haciendo parecer a las supernovas más rojas de lo que en realidad son (de la misma forma que el polvo atmosférico colorea el sol poniente). No observamos ninguna alteración de ese tipo. Y, a menos que mostrara una dispersión regular por todo el espacio, el polvo cósmico tendería a introducir una gran variación en las mediciones, cosa que tampoco vemos.

Se pretende justificar también tal palidez recurriendo al fenómeno de lente gravitatoria. Así se llama la desviación de los rayos de luz al atravesar en su camino las regiones externas de galaxias. El fenómeno de la lente gravitatoria acarrea un aumento ocasional del brillo de fuentes remotas; en la mayoría de los casos, sin embargo, provoca una disminución del brillo. De ese modo podría contribuir a la debilidad de las supernovas distantes. Sin embargo, de los cálculos realizados se desprende que este efecto sólo adquiere significación cuando se trata de fuentes situadas todavía más lejos que las supernovas estudiadas. Podemos, pues, descartar esa explicación.

¿No sería, pensamos, por último, que las supernovas remotas difirieran de las cercanas, quizá porque se formaron de estrellas más jóvenes que albergaban menos elementos pesados que las típicas de galaxias más viejas? No descartamos esta posibilidad, aunque en nuestro análisis procuramos ya tomar en consideración esas diferencias. Estos estudios dan buenos resultados cuando se aplican a galaxias cercanas, cuya edad y morfología varía bastante de una a otra, así como los tipos de supernova observados en ellas.

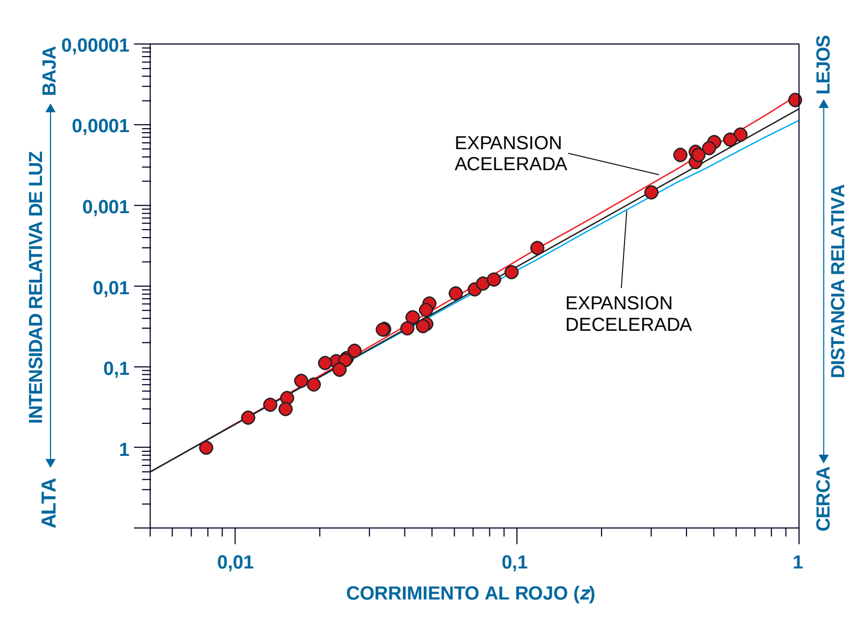
Puesto que ninguno de estos efectos triviales explica las observaciones, nos inclinamos, con otros muchos, a pensar que la debilidad de brillo de las supernovas remotas se debe a la estructura del cosmos. Podrían hallarse involucradas dos propiedades del espacio y del tiempo.

En primer lugar, el espacio puede tener una curvatura negativa. Para entender semejante deformación del espacio, sirvámonos de una analogía bidimensional. Los seres que viven en un mundo plano y de dos dimensiones (como

los personajes de la novela de Edwin A. Abbott *Planilandia*) podrían averiguar que un círculo de radio *r* tiene una circunferencia de exactamente $2\pi r$. Pero si su mundo estuviera ligeramente curvado en forma de silla de montar, tendría una curvatura levemente negativa (*véase* "Inflación en un universo de baja densidad", por Martin A. Bucher y David N. Spergel, *en este mismo número*). Los residentes bidimensionales del país curvado en silla de montar podrían permanecer ignorando esa curvatura hasta que midieran un círculo grande de radio dado y descubrieran que su circunferencia era mayor que $2\pi r$.

La mayoría de los cosmólogos han dado por supuesto, apoyados en razones teóricas, que nuestro espacio tridimensional, igual que *Planilandia*, no está curvado. Pero si poseyera curvatura negativa, la enorme esfera de radiación arrojada por una supernova muy antigua tendría un extensión mayor que la que dispondría en un espacio geométricamente plano, determinando que la fuente apareciera extrañamente debilitada.

Una segunda explicación de la inesperada debilidad de las supernovas remotas lo atribuye a que éstas se hallarían más alejadas que lo que sus corrimientos al rojo sugieren. Con otras palabras, las supernovas situadas a estas enormes distancias mostrarían menos corrimiento al rojo que lo que cabría atribuirles. Para justificar ese menor corrimiento hacia el rojo, los cosmólogos postulan que la expansión del universo procedió con mayor lentitud en el pasado que lo que se supone, con un consiguiente menor desplegamiento global del universo y de la luz que viaja en su seno.




**5. LAS OBSERVACIONES DE SUPERNOVA realizadas por el equipo en el que participan los autores (*puntos rojos*) se desvían ligera, aunque significativamente, del comportamiento esperado: una deceleración bastante rápida (*línea azul*) que ocurriría si el universo fuera "plano" y sin constante cosmológica. Pero las observaciones indican que el universo tiene sólo el 20% de la materia necesaria para hacerlo plano, porque se está decelerando con mayor lentitud que lo predicho (*línea negra*). Las mediciones sugieren que la expansión se está acelerando, debido posiblemente a una constante cosmológica no nula (*línea roja*).**

## La fuerza

¿Qué significa que la expansión cósmica se frene menos deprisa de lo que se venía pensando? Si el universo consta de materia normal, la gravedad ha de frenar sin cesar la expansión. Un frenado escaso, según sugieren las mediciones de las supernovas, nos indica lisa y llanamente que la densidad global de materia distribuida en el universo es baja.

Por muchos supuestos teóricos que socave tal conclusión, no está ayuna de pruebas. En este contexto, algunos astrónomos afirman que ciertas estrellas parecen más viejas que la edad aceptada del universo, obvia contradicción. Pero si el cosmos se expandía antaño con mayor parsimonia, según nos revelan ahora las supernovas, la edad del universo debe revisarse al alza, con lo cual resolveríamos el problema. Los nuevos resultados se hallan también en sintonía con la investigación reciente de la cuantía total de materia alojada en el universo, a través del estudio de los cúmulos galácticos.

¿Qué aclara la nueva determinación de la densidad de materia del universo a propósito de la curvatura del mismo? De acuerdo con los principios de la relatividad general, la curvatura del universo y su deceleración están vinculadas. Parafraseando a John A. Wheeler, la materia le indica al espacio-tiempo cómo curvarse y éste le dice a aquélla cómo moverse. Una baja densidad de materia implica una curvatura negativa y poco frenado. Si el universo está prácticamente vacío, estos dos efectos de atenuación se hallarán cerca de su máximo teórico.

Entendido eso, se explica mejor nuestra sorpresa. Las supernovas que vemos son más débiles que lo predicho incluso para un universo casi vacío (que adquiere máxima curvatura negativa). Tomadas por el valor que indican, nuestras observaciones demandan que la expansión se esté acelerando con el tiempo. Ahora bien, un universo compuesto de manera exclusiva por materia ordinaria no puede crecer así, porque no dejaría de ejercerse la fuerza de atracción gravitatoria. Sin embargo, de acuerdo con la teoría de Einstein, la expansión puede acelerarse si una forma exótica de energía llena el espacio. Esta extraña "energía de vacío" queda incorporada en las ecuaciones de Einstein dentro de la constante cosmológica. A diferencia de las formas comunes de masa y energía, la energía de vacío añade una fuerza repulsiva y puede expandir el universo a velocidades crecientes. Una vez admitimos esta extraordinaria posibilidad, nuestras observaciones encajan perfectamente, incluso suponiendo la geometría plana preferida por los teóricos.

Las pruebas de la presencia de una forma extraña de energía, que ejerce una fuerza gravitatoria repulsiva, constituyen el resultado más desconcertante de cuantos cabía esperar. Y lo es en tal grado, que lo tomamos con reservada cautela. Para disipar dudas contamos con los avances técnicos, tales como los nuevos detectores infrarrojos y el *Telescopio Espacial de Nueva Generación*, que pronto nos permitirán examinar nuestras conclusiones con mayor precisión y fiabilidad. Estos maravillosos instrumentos nos habrán de facilitar la percepción de faros más débiles, que emitieron su destello hace mucho más tiempo en galaxias remotísimas.

# Supernovas y expansión acelerada del universo

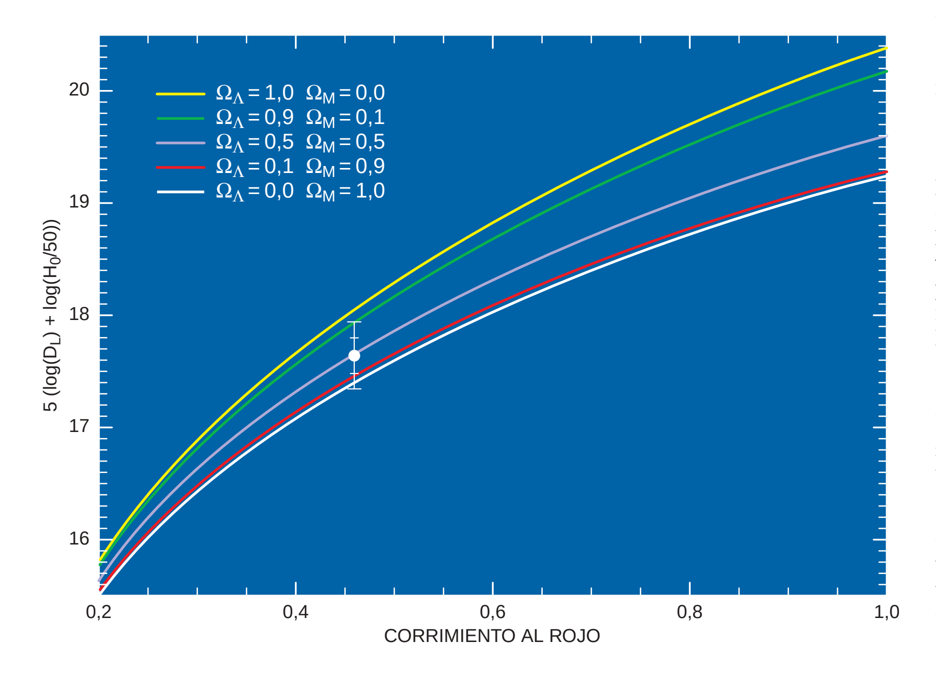
*La exploración del espacio-tiempo por medio de supernovas, además de revelar que el cosmos acelera su expansión, puede arrojar luz sobre la naturaleza del universo y sus constituyentes*

**Pilar Ruiz-Lapuente, Alex G. Kim y Nicholas Walton**

La comunidad científica asistió el año pasado a un doble descubrimiento que entrañaba un cambio radical en la comprensión del cosmos. Por un lado, la aceleración de la velocidad de expansión del universo y, por otro, la necesidad de que una forma exótica de energía sea responsable de dicha aceleración. La energía en cuestión constituiría una parte substancial del contenido del cosmos y contrarrestaría la fuerza ejercida por la atracción gravitatoria en su globalidad.

Dos fueron los equipos de astrofísicos implicados en el nuevo marco. Uno y otro aplicaron los mismos métodos, por lo que no sorprendió que sus conclusiones coincidieran. Cada grupo, sin embargo, trabajó por su cuenta. Sobre la labor realizada por el proyecto Alta-Z ("High-Z Team") se han extendido en el artículo precedente Robert Kirshner y Nicholas Suntzeff, miembros del mismo. Nosotros formamos parte del Proyecto Supernovas y Cosmología ("Supernova Cosmology Project"), en el que intervienen otros astrofísicos de institutos y universidades americanas, europeas y australianas. Desarrollamos parte de la labor en el Observatorio del Roque de los Muchachos, en la isla de La Palma.

El método idóneo para explorar la evolución dinámica del cosmos en el transcurso del tiempo consiste en desentrañar su cambio de tamaño desde la gran explosión inicial ("big bang"), hace miles de millones de




**1. LA OBSERVACION de supernovas a distintos *z* permite discernir entre los diferentes tipos de universo. En el eje de las abscisas tenemos el corrimiento al rojo de la supernova y en el eje de las ordenadas el brillo aparente. En la teoría de la relatividad general, la curvatura del universo depende —lo mismo que la evolución del factor de escala— del contenido de masa-energía del universo. La curvatura, que influye en el brillo, varía en razón del contenido. Puede ser positiva, negativa o nula. La curvatura negativa del universo actúa como una lupa divergente, que rebaja el brillo de un objeto situado a gran distancia de nosotros. La curvatura positiva, por contra, actúa a modo de lente convergente, que refuerza el brillo del objeto lejano. Si juntamos muchas mediciones de brillo y corrimiento al rojo de un objeto uniforme obtenemos una idea de cuál es el contenido de masa-energía del universo, es decir, la densidad de materia expresada en el valor $\Omega_M$ y cuál es el valor de la densidad de energía de vacío $\Omega_\Lambda$.**

años, hasta hoy. Para sondear esa evolución de la talla del universo nos servimos de la luz generada en explosiones de supernovas. Al ir recreando la expansión del cosmos a través de la luz de supernovas remotísimas nos hemos encontrado con que se está acelerando el ritmo. En el curso de esa misma investigación aparece el inventario de las formas de materia y energía que llenan el cosmos. Entre éstas debe existir una forma exótica de energía que empuja al universo a acelerar su expansión.

## Al rojo

En su viaje por un universo en expansión, la luz se propaga con una longitud de onda cada vez más larga, se enrojece. A ese fenómeno se le conoce por "corrimiento hacia el rojo cosmológico" y se simboliza con la letra *z*. Cuando un telescopio recibe luz procedente de una supernova, se recoge una radiación que lleva viajando miles de millones de años, lo que significa que se ha trasladado a longitudes de onda cada vez mayores. Podemos medir el corrimiento hacia el rojo de nuestra supernova a partir de la luz analizada y descompuesta en sus colores, o sea, desde su espectro de luz.

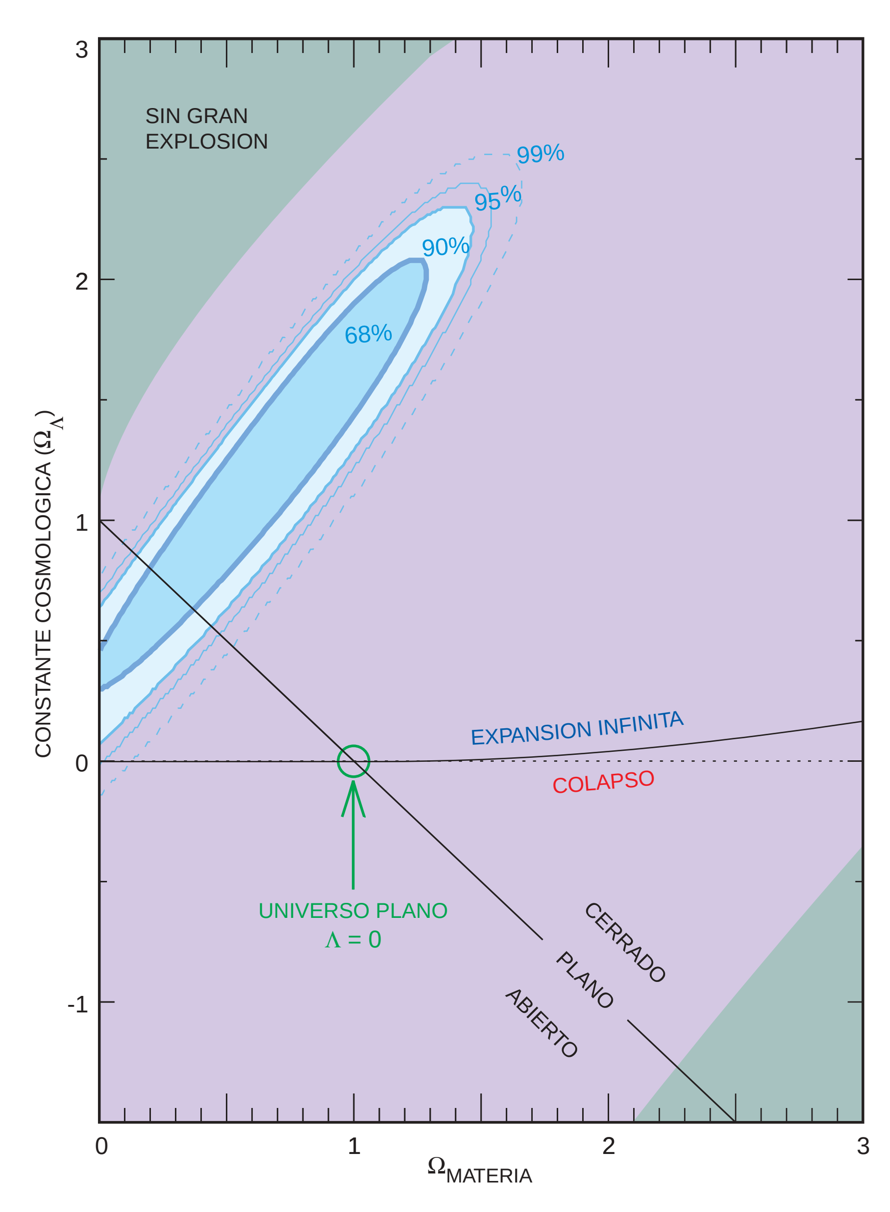
El corrimiento hacia el rojo adquiere pleno sentido en la cosmología de Einstein. En el marco de la teoría general de la relatividad, el corrimiento constituye una medida del factor de escala —o tamaño— del universo en el momento en que se emitió la luz frente al alcanzado en el momento actual. Cuanto mayor es el corrimiento al rojo, más ha cambiado el factor de escala desde el momento de la emisión de la luz y más tiempo ha transcurrido.

De la luz procedente de las supernovas lejanas no sólo nos interesa el corrimiento hacia el rojo. Nos importa y mucho, el brillo aparente. Este flujo viene condicionado por la expansión cósmica. Además de depender del cuadrado de la distancia a la que se encuentra la supernova, el brillo aparente viene afectado por el factor de corrimiento al rojo cosmológico y la dilatación del tiempo en un universo en expansión.

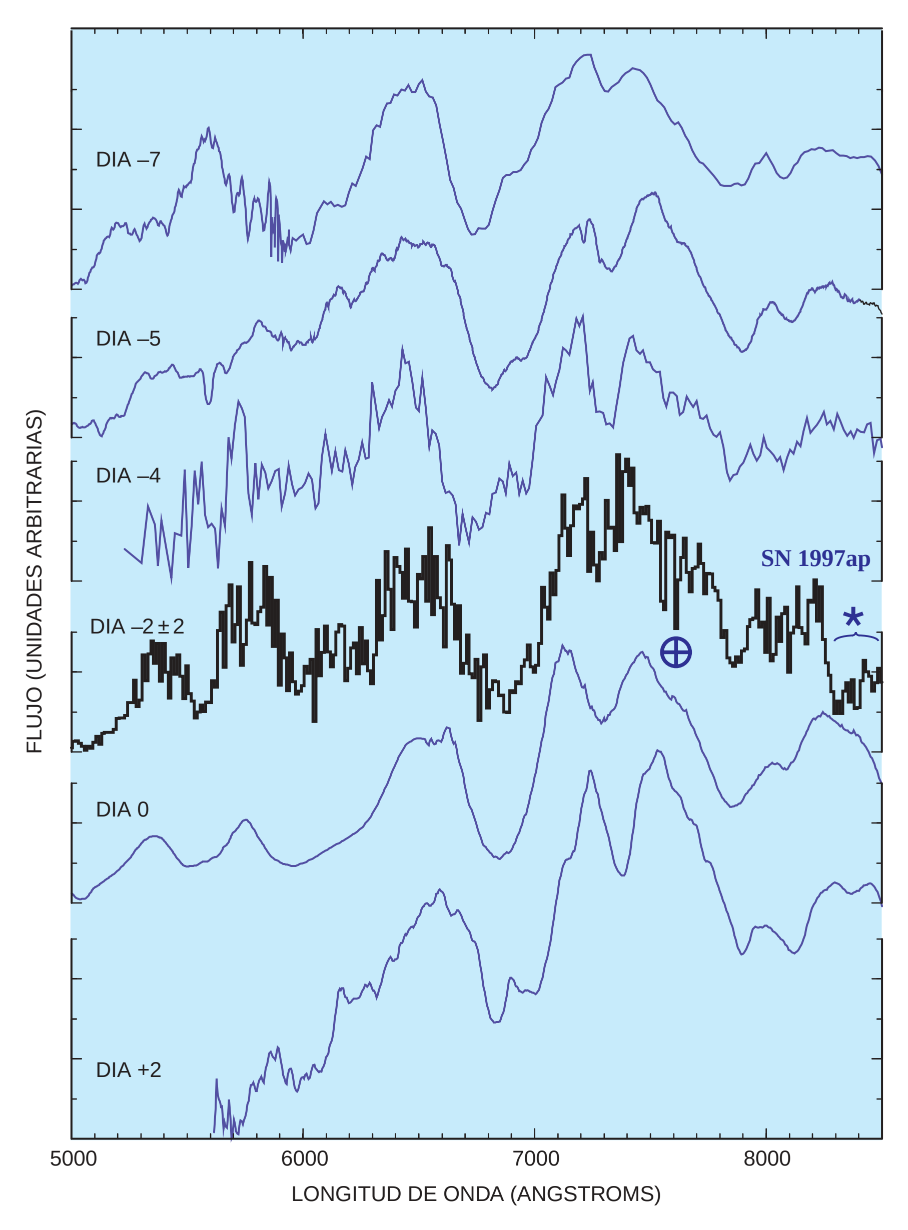
El inverso de la longitud de onda, o frecuencia con que llega la radiación al telescopio, nos da una idea de la energía. La potencia de un pulso de luz emitido en cierto intervalo de tiempo y cierta banda de frecuencia a miles de millones de años luz es más elevada que la potencia de ese pulso en el momento de su recepción en la Tierra. Ello se debe a que la energía recibida ha decrecido, porque la longitud de onda se ha estirado y, en consecuencia, ha disminuido su frecuencia. A su vez, el intervalo de tiempo entre dos pulsos de emisión se ha alargado en su viaje por un universo en expansión. Por eso, cuando medimos, en un intervalo de tiempo fijo, la energía recibida en una banda de frecuencias fija, nos encontramos con que esta cantidad (la potencia recibida) ha decrecido tanto más cuanto mayor ha sido el corrimiento al rojo.

Si dibujamos, siguiendo a Edwin Hubble y Allan G. Sandage, un diagrama donde aparezcan el corrimiento hacia el rojo de un objeto de luminosidad intrínseca conocida y su brillo aparente, observaremos una relación entre el cambio de factor de escala del universo y la distancia. Podemos poner muchos valores en el diagrama, para conocer de qué modo ha cambiado el factor de escala del universo con el tiempo. Ello nos indica el tipo de universo en el que estamos.

La idea de recurrir a las supernovas es bastante reciente; se tomó tras años de estudio por teóricos y especialistas observacionales. Bus-




**2. RESULTADOS DEL PROYECTO SUPERNOVAS Y COSMOLOGIA para el valor de $\Omega_M$ y de la contribución de la constante cosmológica $\Omega_\Lambda$. De los datos se infiere que la densidad de materia del universo está por debajo del valor crítico para "cerrarlo" y que la constante cosmológica tiene un valor no nulo mayor que cero.**




**3. ESPECTRO DE RADIACION de una supernova termonuclear descubierta a unos 7000 millones de años-luz comparada con una muestra más cercana a tan sólo decenas de millones de años-luz. Estos objetos estelares tienen las mismas características de luminosidad en el universo joven, de hace miles de millones de años, que en el universo actual.**

cábase, en efecto, un objeto cuyas propiedades de luminosidad fueran universales —independientes del tiempo y de la región del espacio donde se producen—, homogéneas —con poca dispersión en su brillo intrínseco o con propiedades de brillo intrínseco que obedezcan a una ley o relación bien conocida— y observables hasta muy altos corrimientos al rojo.

Las explosiones termonucleares de enanas blancas, conocidas como supernovas de tipo Ia o supernovas termonucleares, resultaron ser las que mejor se adecuaban a los propósitos de la investigación. Así, en lo concerniente a la homogeneidad espaciotemporal hemos demostrado que las propiedades de brillo de este objeto se definen con una precisión del 10 %, independientemente del entorno.

Aunque no todas las supernovas termonucleares son iguales, pues unas brillan más que otras, todas siguen un mismo patrón en su luminosidad. La correlación funciona de forma tan precisa que, conociendo la anchura de la curva de luz de la supernova, sabemos su luminosidad intrínseca. Tenemos una especie de gigantescas bombillas en las que con sólo leer la etiqueta conocemos los watts de potencia que vamos a tener.

## Expansión acelerada

El contenido de materia y energía es el responsable del cambio en la velocidad de expansión del universo. En relatividad einsteiniana, materia y energía influyen en la dinámica del espacio-tiempo en un plano de igualdad.

El primer intento de medir el valor actual de la velocidad de expansión del universo lo realizó Hubble. En su honor se denominó parámetro de Hubble la velocidad con que se expande el universo. Su valor actual se conoce por constante de Hubble ($H_0$). Un segundo factor pone de manifiesto el cambio de ritmo de la velocidad de expansión en virtud del contenido de masa-energía del cosmos. Nos referimos al parámetro de deceleración, cuyo valor actual se representa por $q_0$. El parámetro $q_0$ es positivo si el universo se decelera y negativo, si se acelera. La dinámica de la aceleración o desaceleración es, en última instancia, función de la densidad total del universo y del comportamiento dinámico de ese contenido, de si contribuirá a la expansión o la frenará.

Esto último viene dado por la relación entre presión ($p$) y densidad ($\rho$) de cada componente de materia-energía del cosmos, lo que se conoce como ecuación de estado de cada componente. Si el universo está dominado por una componente en forma de materia o energía de presión negativa hasta el punto de que $q_0$ adquiera valores negativos, la expansión se acelera. En general, las formas conocidas de materia y energía tienen presión positiva y frenan la expansión.

Si la densidad de materia del cosmos es muy grande, el frenado terminará por invertir el ritmo de expansión y cambiarla de signo; el universo retrocedería en sus dimensiones hasta desplomarse en un colapso final. Esto ocurre cuando $\Omega_0$, parámetro de densidad del universo, es mayor que 1, y no domina la componente de presión negativa.

Para entender el papel de los componentes que contribuyen a frenar o acelerar el cosmos, fijémonos en una forma común de energía, la luz o radiación electromagnética. Si tenemos un volumen lleno de radiación, ésta ejerce una presión positiva sobre las paredes que delimitan el volumen. La presión aumenta con la densidad de ra-

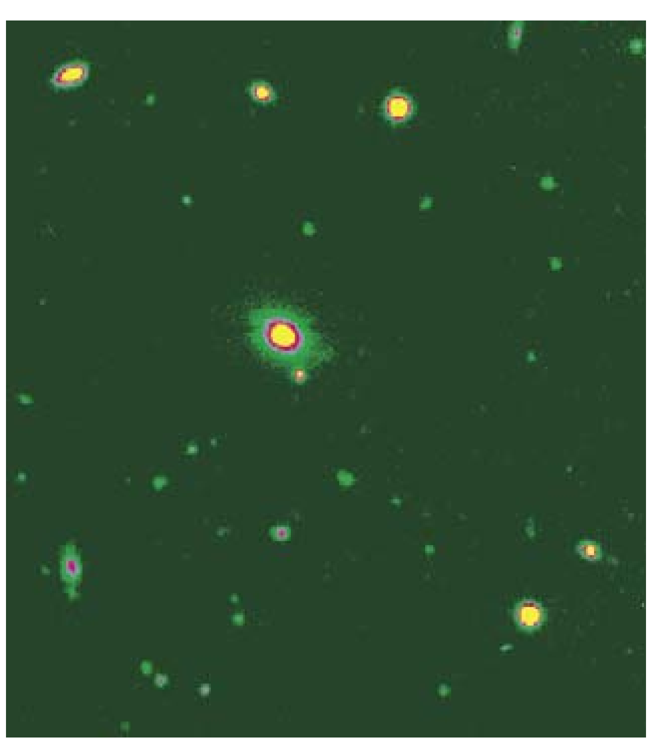
## El Roque de los Muchachos y el Proyecto Supernovas y Cosmología

Dentro del Proyecto Supernovas y Cosmología corresponde al Observatorio del Roque de los Muchachos, en la isla de La Palma, acometer programas de descubrimiento y obtención de curvas de luz y espectros de supernovas que distan de nosotros miles de millones de años luz. Sus datos se combinan con los de otros observatorios agregados al proyecto. Sus resultados deben compararse y completarse con las mediciones y análisis efectuados desde el Observatorio Europeo del Hemisferio Sur en Chile, el Kitt Peak de Arizona, el Keck de Hawai y el Telescopio Espacial Hubble.

Nos ocupamos, en particular, de los posibles efectos sistemáticos que puedan introducirse en el método. Con ese fin vamos acopiando muestras de supernovas a distintos corrimientos al rojo ($z$), en busca de diferencias sistemáticas entre las muestras más lejanas y las más cercanas. Se trata de observar los mismos campos de galaxias con una diferencia de unas tres semanas, identificar las supernovas que se producen y seguir con rigor la curva de luz.

El descubrimiento y el seguimiento de la curva de luz se efectúa en el telescopio Isaac Newton de 2,5 metros, que está provisto de una cámara de amplio campo y detectores de gran eficiencia. La identificación espectral de las supernovas se realiza en el telescopio de 4,2 m William Herschel. Para ello necesitamos medir con gran precisión la posición de las supernovas en el cielo en nuestras imágenes del telescopio Isaac Newton. Por ser objetos muy débiles, no se perciben en la cámara de ayuda instalada en el William Herschel. Ahora bien, el telescopio puede apuntarse con exquisita precisión; hasta el segundo de arco, que es la precisión necesaria para situar la rendija de nuestro espectrógrafo de modo que capte la luz del objeto y descomponga la luz que se recibe en sus longitudes de onda. Para obtener un espectro necesitamos sumar luz durante una hora, o incluso más, dependiendo de la distancia.

En el telescopio Isaac Newton y en el telescopio Nórdico (de 2,6 m) situado en el Observatorio del Roque seguimos las supernovas de más alto $z$ que nuestros colegas de otros institutos han descubierto. Con la coordinación de estas labores y el análisis detallado de los datos es posible llegar a resultados óptimos.



**2. IMAGEN DE UNA SUPERNOVA (*indicada por la flecha*) en una galaxia elíptica a unos miles de millones de años-luz. Imagen obtenida con el telescopio Isaac Newton.**



**1. TELESCOPIO ISAAC NEWTON del Observatorio del Roque de los Muchachos. Con él descubrimos supernovas a alto *z*.**

diación, en una proporción de $p = {}^1/_3\rho$. El índice que caracteriza esta forma de energía ($w = {}^1/_3$) es positivo.

A lo largo del tiempo, la evolución de esta energía de radiación difiere de la evolución seguida por la materia. Conforme se expande el universo, la densidad de materia decrece de forma inversa al aumento de volumen; evoluciona de acuerdo con una razón o ley de $t^{-3}$. Cuando la escala del universo crece proporcionalmente al tiempo, la densidad de radiación decrece con una potencia mayor, $t^{-4}$; llegará así un momento en que la densidad de radiación pierda importancia, comparada con la materia, en el desarrollo dinámico del universo.

En el universo cercano a la gran explosión, la densidad de radiación dominaba sobre la densidad de materia. Desde entonces empezó a decrecer deprisa la densidad de radiación; ahora, la componente de radiación ejerce un efecto despreciable sobre la dinámica del cosmos. A diferencia de lo que acontecía en los momentos subsiguientes a la gran explosión —con predominio de la radiación— nos encontramos ya en la era en que el papel principal de la evolución del cosmos lo representa la materia.

La materia que interviene en la evolución dinámica del universo abarca la materia ordinaria y la materia oscura. Conocemos de ésta

sus efectos en el movimiento de las estrellas, pero ignoramos de qué partículas consta. En nuestro proyecto medimos los valores globales de la densidad de materia (toda la que ejerce un efecto dinámico, no sólo la hecha de protones y neutrones) y de la densidad de energía del universo que mejor se ajustan a las observaciones.

Durante el año 1997, conforme íbamos recogiendo y analizando datos de supernovas nos dimos cuenta de un primer resultado con importantes implicaciones: la baja densidad del universo, muy inferior al valor crítico que determinaría la inversión de su evolución hasta acabar en un gran colapso ("big crunch"). Podemos descartar esa evolución futura en un universo en que las galaxias se acercarían unas a otras chocando a un ritmo cada vez mayor con el encogimiento del cosmos.

Nuestro universo se irá apagando e hinchándose indefinidamente. El destino del universo es el de una expansión que diluye la materia en un volumen cada vez mayor y separa a distancias infinitas los restos de las galaxias, cada vez menos brillantes.

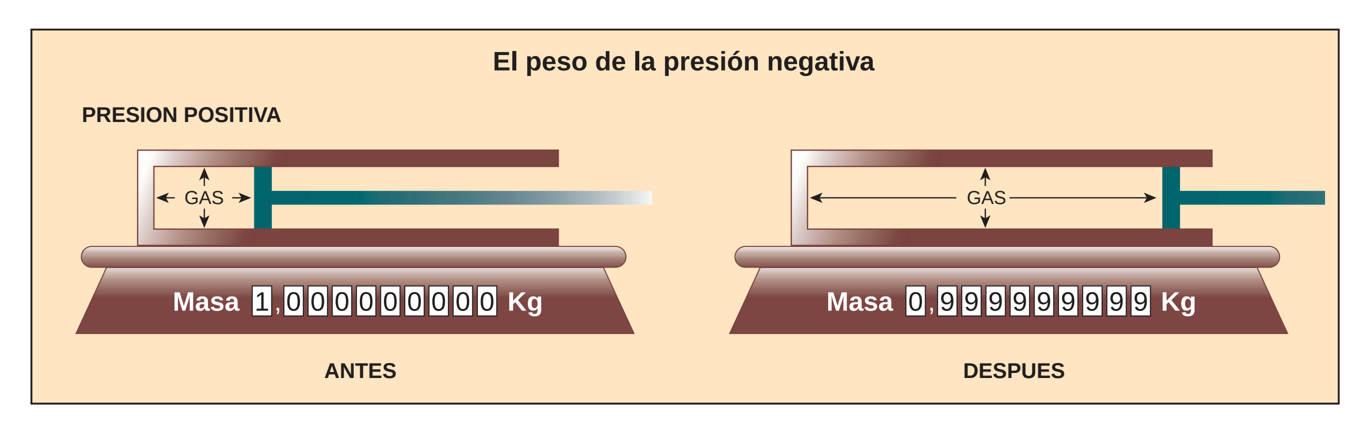
Cuando afirmamos que la densidad de materia es baja nos referimos a que el parámetro de densidad de materia $\Omega_m$ es menor que 1. Nos dimos cuenta de esa baja densidad trabajando con una muestra de supernovas cuya mayoría no presentaba un valor de z muy alto. A ese resultado le acompañó otro. Conforme observábamos y analizábamos más supernovas, y más alejadas, se nos reveló algo inesperado. La exploración del cambio de escala del universo, merced a la investigación con supernovas, nos descubría que en el pasado el proceso de agrandamiento avanzó más lento. O lo que era lo mismo: el universo, en vez de disminuir su velocidad de expansión, la aumentaba.

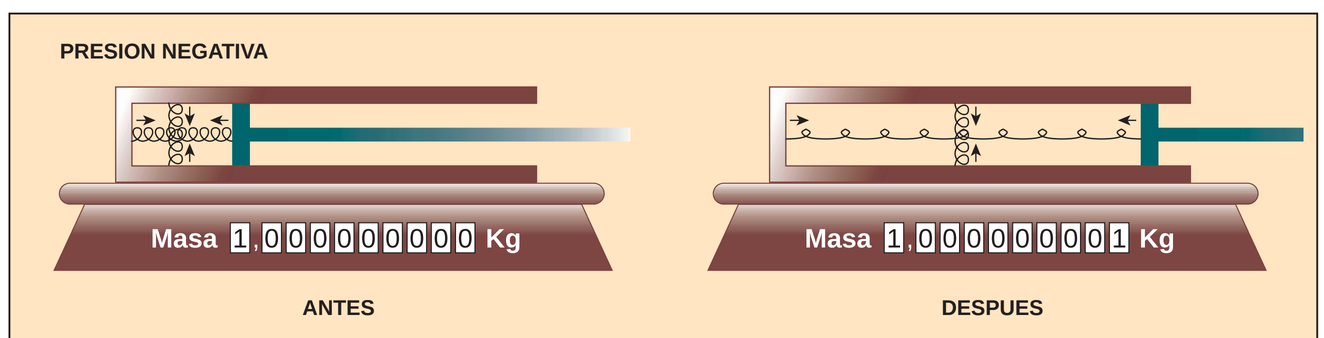
El universo que nos ha dibujado la observación de supernovas muy lejanas está muy vacío y dominado por una componente de energía de presión negativa, ya que sólo así se explica la aceleración de la expansión. Su identificación, objeto de debate todavía, apunta hacia la constante cosmológica, o energía de vacío $\Lambda$ (densidad dada en $\Omega_\Lambda$, parámetro de densidad de energía de vacío). La energía asociada a la constante cosmológica llena el espacio con densidad constante; hace poco podría haber pasado a dominar sobre la materia en la evolución dinámica del universo, de la misma forma que la materia empezó a dominar sobre la radiación en un momento primitivo del cosmos.

## $\Lambda$, la constante cosmológica

Semejante descubrimiento de la existencia de una forma de energía que actúa como una "energía de vacío" ha sido recibido con sumo interés por los físicos que estudian los primeros instantes del universo. De acuerdo con el nuevo modelo, la expansión se está acelerando y el pistón que la empuja es una componente que además parece dominar sobre la materia. En efecto, en el balance de energía-materia registrado, un 30 % lo proporciona la materia y un 70 % corresponde a la energía de vacío.

### El peso de la presión negativa




**4. EXPERIMENTO IDEAL. La masa-energía en un pistón llenado con una substancia de presión positiva (*arriba*) y negativa (*abajo*) cambia en virtud de la expansión del émbolo. Para que la presión fuera del pistón no desempeñe ninguna función, demos por supuesto que el experimento se hace en el vacío. Arriba pesamos un pistón lleno con gas comprimido, lo expandimos y volvemos a pesar. En el panel inferior, hacemos lo mismo con una substancia de presión negativa, representada por el muelle estirado. La balanza es sensible a pequeños cambios en la masa-energía (en relatividad estas dos cantidades son equivalentes). La substancia de presión positiva, al expandirse en un volumen mayor, relaja su densidad y energía interna. Pesa menos en la posición final que en la inicial. La substancia de presión negativa tal como el muelle dentro del pistón (o, en nuestro universo, la energía de vacío o constante cosmológica) aumenta su energía interna al expandirse y pesa más en nuestra balanza.**

## ¿No hay efectos de evolución?

El fundamento del método escogido para estudiar el ritmo de expansión del universo fallaría si las supernovas de alto corrimiento hacia el rojo siguieran relaciones diferentes de las observadas en las supernovas de bajo corrimiento al rojo. Las estrellas que dan lugar a enanas blancas y explotan en supernovas termonucleares tardan miles de millones de años en recorrer ese trayecto. El entorno en que nació una estrella hace 10.000 millones de años (la edad de las supernovas más lejanas) y el medio que acunó a otro de sólo decenas de millones de años (supernovas de la colección más cercana) diferían notablemente. Sin embargo, son muy parecidas al observador. ¿Por qué?

A lo largo de la evolución que da lugar a la enana blanca, la estrella original, con una masa de entre 3 y 10 veces la solar, ha ido perdiendo sus capas externas y comprimiéndose bajo el efecto de la gravedad. Al final, lo que queda de la estrella es un núcleo comprimido hasta una densidad que multiplica millones de veces la de la materia ordinaria. Este núcleo estelar contiene carbono y oxígeno; los electrones están tan confinados, que son compartidos entre los distintos átomos y se comportan como un único gas. La presión de este gas de electrones mantiene a la estrella frente a la contracción gravitatoria. Pero no nos hallamos ante un gas ideal que incremente la presión en respuesta a la mayor temperatura o a una contracción. La presión no aumenta con la temperatura; la estrella no puede relajar su estructura expandiéndose.

Si la enana blanca gana masa al atraer materia de una estrella compañera, la densidad y temperatura de su interior pueden seguir creciendo sin que se expanda el material del interior. Cuando la temperatura adquiere cierto valor, la estrella estalla al desencadenarse en su interior reacciones nucleares que queman el carbono y el oxígeno y originan elementos pesados.

La explosión termonuclear puede ser más o menos energética y sintetizar más o menos elementos pesados. De éstos algunos son radiactivos y emiten radiación gamma en su desintegración; son los responsables de la luminosidad de las supernovas.

Las explosiones que han quemado más carbono y oxígeno y forman más elementos pesados engendran supernovas más brillantes; las explosiones que han quemado poco material dan supernovas menos brillantes. La evolución de la luminosidad en el tiempo que experimentan unas y otras, muy regular, no depende del entorno en que se ha producido la explosión.

No queda ningún residuo estelar. El material quemado se expande a enormes velocidades y la radiación que lo atraviesa experimenta interacciones con los átomos del material en expansión dando lugar a un espectro característico que va cambiando a medida que pasa el tiempo. Tan regular es también el espectro, que al medirlo vemos a qué fase de la explosión corresponde.

He ahí la respuesta al por qué de arriba: la estrella que acaba explotando como supernova termonuclear, desprovista de sus ornamentos externos, quedó reducida a un núcleo de estructura definida sólo por su masa.

No es el único ejemplo de estrella cuya estructura queda prácticamente definida por un solo parámetro. Las estructuras de las estrellas de neutrones y los agujeros negros también quedan definidas por la masa. Pero, a diferencia de las enanas blancas, las estrellas de neutrones y los agujeros negros no explotan. Y por tanto, no proporcionan el brillo calibrado que nos proporcionan las enanas blancas.

A esa energía de vacío algunos han empezado a llamarla energía oscura, en paralelismo con la materia oscura cuya naturaleza desconocemos, si bien detectamos sus efectos dinámicos en las galaxias.

Podemos imaginarnos la constante cosmológica como una forma exótica de materia cuya densidad no cambia al expandirse o contraerse el volumen que la confina. No es intuitiva la idea de una densidad independiente del volumen; antes bien, la experiencia diaria nos muestra que la materia se diluye, se torna menos densa, al aumentar el volumen. La posibilidad de que una substancia mantenga su densidad al expandirse es una consecuencia directa de la relación de Einstein, $E = mc^2$. Al ser equivalentes materia y energía, si expandimos una substancia elevando su energía interna podremos conseguir que su densidad no decrezca.

Ilustraremos ese efecto con un sencillo experimento ideal. Sea un pistón que confina substancias con presión positiva y negativa, respectivamente. Por substancia de presión positiva tomaremos un gas intensamente comprimido; por substancia de presión negativa, un muelle sujeto por los cuatro lados a las paredes del pistón y que se estira al expandirse el volumen del mismo. El muelle ejerce una presión sobre las paredes del volumen negativa, en la dirección que tiende a reducir el volumen confinado por el pistón.

En el experimento ideal se utiliza una balanza capaz de medir masa-energía, para detectar el cambio operado en la masa y causado por la expansión del pistón. En el caso del gas com-

PILAR RUIZ-LAPUENTE, ALEX G. KIM y NICHOLAS WALTON comparten un mismo interés por las explosiones de supernova y colaboran en el Proyecto Supernovas y Cosmología (“Supernova Cosmology Project”). Ruiz-Lapuente, profesora de la Universidad de Barcelona, estudia las últimas etapas de la evolución estelar. Ha sido investigadora contratada por la Universidad de Harvard e invitada por el Instituto Max-Planck de Astrofísica de Garching. Kim trabaja en el Laboratorio de Física Corpuscular y Cosmología del Colegio de Francia en París e investiga asimismo sobre la naturaleza de la materia oscura. Walton, astrofísico especialista en instrumentación, desarrolla su investigación en el grupo Isaac Newton del observatorio del Roque de los Muchachos. Además del final explosivo de estrellas le interesa la formación de nebulosas planetarias.

primido, con presión positiva, el gas pierde energía interna al expandirse el pistón. Debido a la relación $E = mc^2$ esta disminución en energía interna se traduce en una disminución de la masa, medida por la balanza de fina precisión. Por otra parte, en el caso de la substancia de presión negativa, representada por el muelle, se debe realizar trabajo para estirar el resorte, trabajo que aumenta la energía interna del muelle y, por tanto, su masa, según detecta la balanza.

Para muelles hechos de materia ordinaria el incremento de energía interna y masa es minúsculo. Mas, para la constante cosmológica, la presión negativa es enorme y el aumento de masa es proporcional al volumen, de forma que la densidad no cambia al crecer el volumen.

La ecuación de estado para Λ, la energía de vacío, es $p = -\rho$. Para la materia ordinaria (salvo en el caso en que se mueve a velocidades cercanas a la de la luz) la presión es insignificante. Si tenemos materia y energía de vacío como componentes del cosmos, la deceleración viene dada por $q_0 = {}^1/_2\ \Omega_m - \Omega_\Lambda$. En estos términos hemos presentado nuestros resultados.

Pero los resultados de la aceleración de la expansión son también compatibles con la presencia de una forma de energía distinta de la energía de vacío Λ, cuya ecuación de estado sea $p = w\rho$, con $w$ negativo aunque distinto de –1. En esta hipótesis, la energía de vacío podría ser cero.

Para explicar la aceleración habría que buscar, pues, otra componente energética de naturaleza diferente. Algunos optan por esta opción, sabedores de que, en el universo primitivo, la energía de vacío es la energía asociada a la pérdida de simetrías entre las distintas fuerzas fundamentales. En ese contexto, valores como el deducido del diagrama de supernovas son difíciles de explicar: los valores esperados son o bien cero o valores enormemente altos.

No faltan quienes prefieren a la constante cosmológica un tipo de campo escalar que tenga ecuación de estado variable con el tiempo. Este campo habría ido evolucionando desde la gran explosión y ahora estaría en un punto en que su energía asociada da el 70 % de la densidad de masa-energía del universo. La diferencia de esta componente en relación a la constante cosmológica estriba en la densidad de energía, que es variable con el tiempo, mientras que en el caso de la constante cosmológica no lo es. El índice de la ecuación de estado cambia y no llega a ser –1 en la actualidad (–1 es lo que caracteriza a la constante cosmológica).

Otro candidato a componente con presión negativa lo constituyen los defectos topológicos en la textura del espacio-tiempo que ejercen presión negativa y pueden llevar asociadas una densidad de energía considerable.

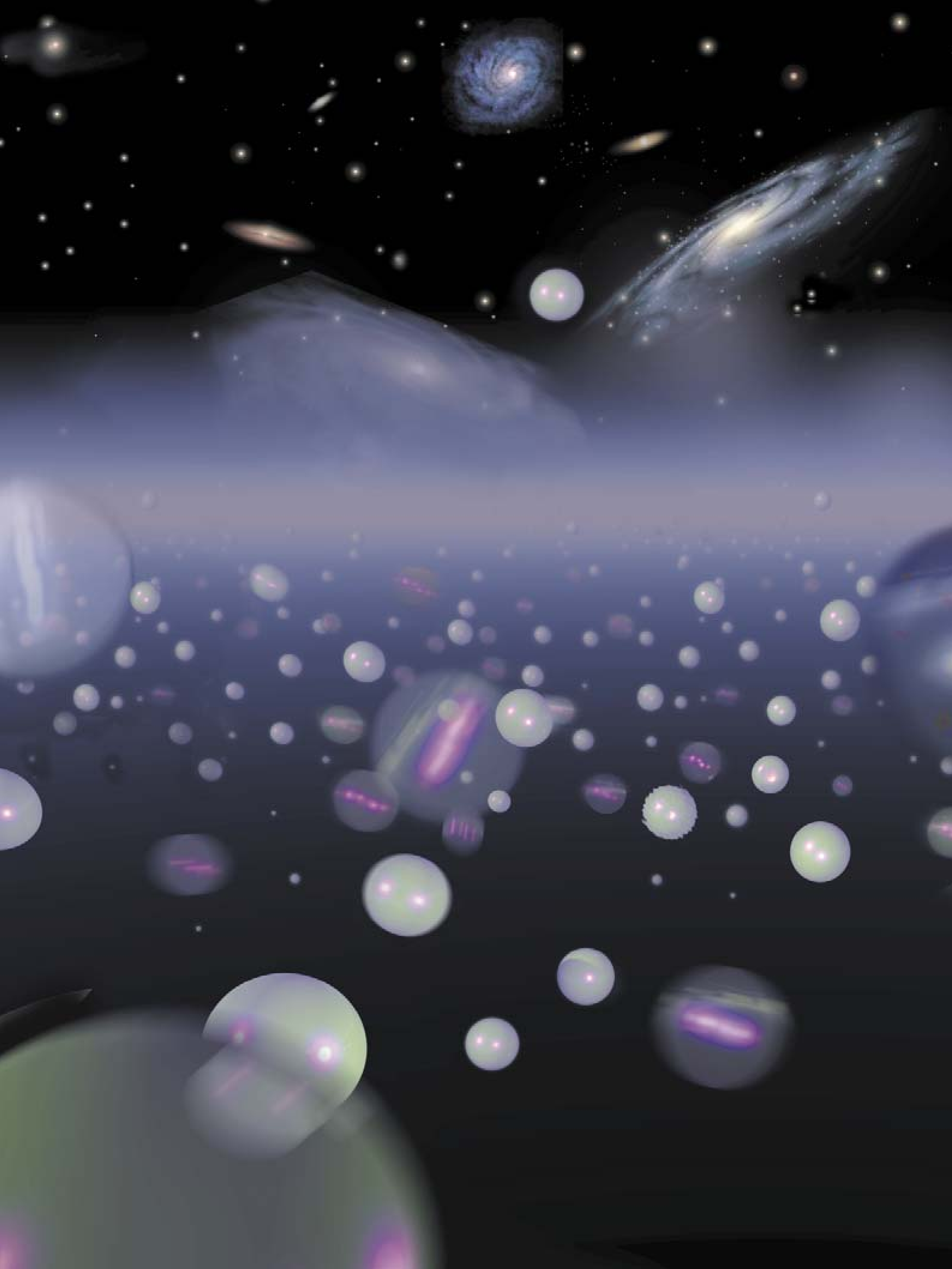
Lo mismo el equipo High-Z que el nuestro proseguimos estudiando la naturaleza del componente de presión negativa. De momento, los resultados provisionales excluyen algunos candidatos a energía oscura; quedarían fuera algunos tipos de defectos topológicos y campos escalares, mientras que Λ, la constante cosmológica, parece la opción más favorable. En los próximos años, prosiguiendo este trabajo, esperamos poder trazar la historia de la aceleración de la expansión del universo y ayudar a discernir cuál es la característica auténtica de lo que llena el cosmos.

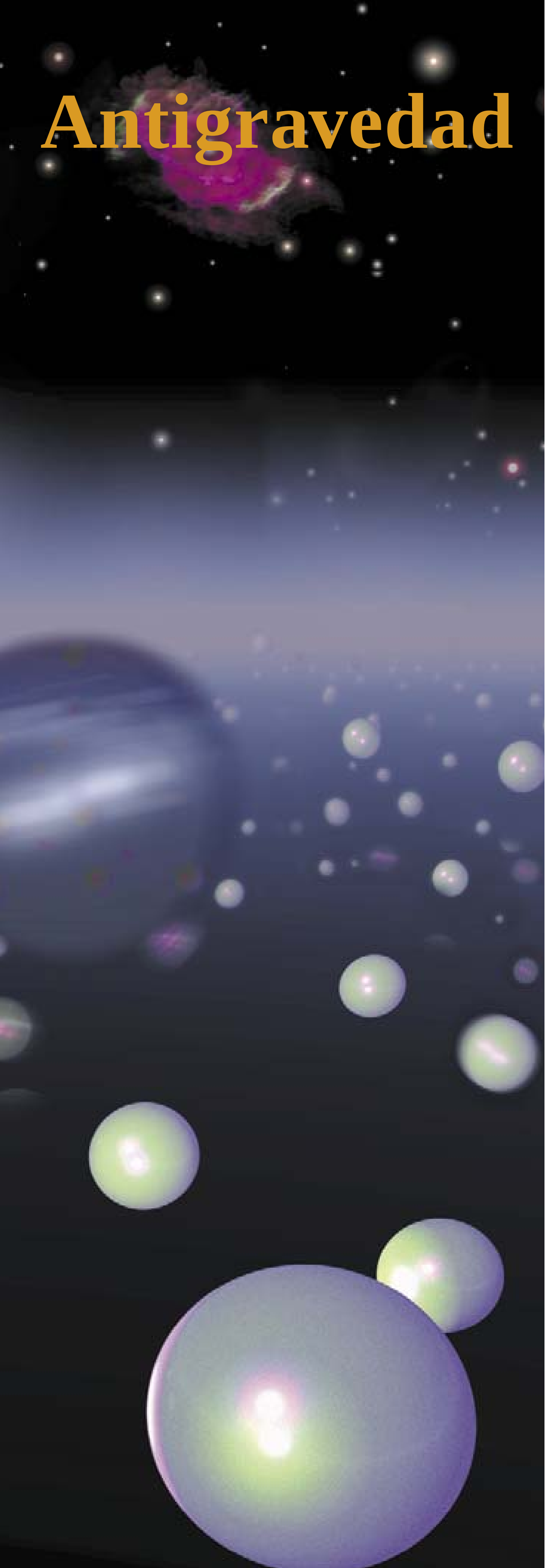
BIBLIOGRAFIA COMPLEMENTARIA

THERMONUCLEAR SUPERNOVAE, Dirigido por P. Ruiz-Lapuente, R. Canal y J. Isern. Kluwer Academic Publisher 1997.

MEASUREMENTS OF $\Omega_M$ AND $\Omega_\Lambda$ FROM 42 HIGH-REDSHIFT SUPERNOVAE. S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaeffer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg y W. J. Couch. (The Supernova Cosmology Project) en *Astrophysical Journal* (en prensa). Preprint en xxx.lanl.gov/abs/astro-ph/ 9812133.

SUPERNOVA LIMITS ON THE COSMIC EQUATION OF STATE. P. M. Garnavich, S. Jha, P. Challis, A. Clocchiatti, A. Diercks, A. V. Filippenko, R. L. Gilliland, C. J. Hogan, R. P. Kirshner, B. Leibundgut, M. M. Phillips, D. Reiss, A. G. Riess, B. P. Schmidt, R. A. Schommer, R. C. Smith, J. Spyromilio, C. Stubbs, N. B. Suntzeff, J. Tonry y S. M. Carroll en *Astrophysical Journal* (en prensa). Preprint en xxx.lanl.gov/abs/astro-ph/9806396.





# Antigravedad cosmológica

*La constante cosmológica, una peculiar forma de energía inherente al espacio mismo, se ha convertido en una de las dos explicaciones plausibles de la expansión acelerada*

**Lawrence M. Krauss**

Decía en 1946 George Orwell, novelista y crítico social, que "para ver lo que se tiene ante las narices hay que luchar sin parar". Estas palabras vienen muy a cuento para describir la cosmología moderna. El universo nos rodea —somos parte de él— y, sin embargo, a veces hay que mirar lo que pasa lejos para conocer los procesos que rigen nuestra vida. Y aunque creemos que los principios que gobiernan la naturaleza son simples, el desentrañarlos ya es harina de otro costal. El firmamento da pistas sutiles. El dicho de Orwell es doblemente cierto para los cosmólogos que han de interpretar las recientes observaciones de estrellas que estallaron a cientos de millones de años de luz de distancia. En contra de lo esperado, la expansión del universo, en vez de frenar su velocidad, la acelera.

Se sabe desde 1929, por lo menos, que el universo visible se expande. Ese año Edwin P. Hubble mostró que las galaxias lejanas se van separando como si el cosmos entero estuviese hinchándose. Semejante alejamiento actúa contra el efecto opuesto ejercido por la gravedad conjunta de cúmulos galácticos y de todos los planetas, estrellas, gases y polvo que contienen. Hasta la minúscula atracción gravitatoria de un clip retarda un poco la expansión cósmica. Hace diez años la concordancia de la teoría con la observación daba a entender que había suficientes clips, suficiente materia en el universo para detener casi —pero sólo casi— la expansión. Por decirlo en la forma geométrica que Einstein promovió, parecía que el universo era "plano".

**1. EL LLAMADO ESPACIO VACIO está repleto de partículas elementales que surgen y desaparecen muy deprisa. No podemos detectarlas directamente. Su presencia obedece a un principio básico de la combinación de la mecánica cuántica con la relatividad especial: nada es exacto, ni siquiera la nada misma. La energía conjunta de estas partículas "virtuales", como otras formas de energía, podría ejercer una fuerza gravitatoria, de atracción o repulsión en razón de principios físicos mal conocidos por ahora. A escala macroscópica esa energía podría actuar como la constante cosmológica propuesta por Albert Einstein.**

## Tipos de materia

| Tipos de materia | Composición probable | Indicio principal | Contribución aproximada a Ω |
|---|---|---|---|
| Materia visible | Materia ordinaria (compuesta sobre todo de protones y neutrones) que forma las estrellas, el polvo y el gas | Las observaciones telescópicas | 0,01 |
| Materia oscura bariónica | Materia ordinaria demasiado oscura para dejarse ver, quizás enanas marrones o negras (objetos compactos de gran masa del halo, los MACHO) | Los cálculos de la nucleosíntesis en la gran explosión (big bang) y la abundancia del deuterio observada | 0,05 |
| Materia oscura no bariónica | Partículas exóticas como los "axiones", los neutrinos con masa o las partículas de gran masa que interactúan débilmente (los WIMP) | La gravedad de la masa visible no basta para explicar las velocidades orbitales de las estrellas dentro de las galaxias y de las galaxias en los cúmulos | 0,3 |
| "Materia oscura" cosmológica | La constante cosmológica (la energía del espacio vacío) | El fondo de microondas da a entender que el cosmos es plano, pero no hay suficiente materia bariónica o no bariónica para que lo sea | 0,6 |

**2. EN EL UNIVERSO SE ALBERGAN miles y miles de millones de galaxias, cada una de las cuales contiene un número no menos abrumador de estrellas. Y, sin embargo, parece que en su mayor parte es "materia oscura". La constante cosmológica, si se confirma su existencia, actuaría a escala cósmica como una forma aún más exótica de materia oscura. La magnitud omega, Ω, es el cociente de la densidad de materia o energía y de la densidad necesaria para que el universo sea plano.**

La del universo plano es una geometría intermedia entre otras dos posibles, la "abierta" y la "cerrada". En un cosmos donde la materia batalla con el impulso expansivo de la gran explosión ("big bang"), la geometría abierta es la victoria de la expansión: ésta proseguiría sin límite temporal. En el universo cerrado, ganaría la gravedad; el mundo acabaría por desplomarse sobre sí mismo de nuevo y terminaría en un ardiente gran implosión ("big crunch"). Los órdenes de cosas abierto, cerrado y plano vienen a ser como lanzar un cohete más deprisa, más despacio o exactamente a la velocidad de escape de la Tierra, la necesaria para superar la atracción gravitatoria del planeta.

Al decir de uno de los postulados clave de la teoría inflacionaria, vivimos en un universo plano, donde existe un perfecto equilibrio de fuerzas. Esa misma teoría, para resolver varias paradojas de la formulación corriente de la gran explosión, dicta que muy al principio del universo hubo un período de expansión rápida. Aunque el contenido visible del cosmos no basta para pensar en un universo plano, la dinámica celeste indica que hay mucha más materia que la percibida. La mayor parte de la encerrada en las galaxias y agrupaciones de galaxias tiene que ser invisible para los telescopios. Hace más de diez años denominé "quintaesencia" a esa materia a la que llaman oscura; tomé la palabra de Aristóteles, que así aludía al éter, elemento invisible que bañaba el espacio [*véase* "Materia oscura en el universo", de Lawrence M. Krauss; INVESTIGACIÓN Y CIENCIA, febrero de 1987].

Pero ahora contamos con un mazo de pruebas que van más allá. Ni siquiera la materia ocultada basta para producir un universo plano. Quizás el mundo no sea plano, sino abierto, en cuyo caso habrá que modificar, o desechar, la teoría inflacionaria. O tal vez sí lo sea, en cuyo caso no tendrá sus constituyentes principales en la materia visible, la oscura o la radiación, sino en una forma aún más etérea de energía que poblaría el espacio vacío.

La idea de esa forma de energía tiene una historia larga y agitada. Empieza cuando Einstein completó su teoría general de la relatividad, más de diez años antes de que Hubble enseñase convincentemente que el universo se expande. La relatividad entrelazaba espacio, tiempo y materia. Prometía así lo que hasta entonces había sido imposible: el conocimiento científico no sólo de la dinámica de los objetos del universo, sino también de la dinámica del mismo cosmos. Sólo había un problema. Al contrario que las demás fuerzas fundamentales a las que se halla sujeta la materia, la gravedad es siempre atractiva: sólo tira de, nunca empuja. La inexorable atracción gravitatoria de la materia podría hacer que el universo acabase por desplomarse. Ante semejante conclusión, Einstein, que imaginaba un universo estático y estable, añadió un término a sus ecuaciones, un "término cosmológico", que lo estabilizaba con una nueva fuerza de largo alcance en el espacio. Si su valor era positivo, sería una fuerza repulsiva, una especie de antigravedad que impediría que el universo se derrumbase bajo su propio peso.

Pero, ay, Einstein abandonaría cinco años después este apaño, del que dijo que había sido su "mayor pifia". La estabilidad que ofrecía el término era ilusoria, y aun contaba más en su contra el respaldo observacional creciente a la tesis de la expansión. En 1923 Einstein le escribía a Hermann Weyl que "si el mundo no es cuasiestático, ¡fuera el término cosmológico!". Como el éter antes, parecía que el término acabaría en la papelera.




**3. ESTA CARTA DE EINSTEIN al matemático Hermann Weyl reconoce que un universo de tamaño inmutable podría ser propenso a la expansión o la contracción: "En el universo de De Sitter dos puntos de fluido inestables distintos se separan aceleradamente. Si el mundo no es cuasiestático, ¡fuera el término cosmológico!"**

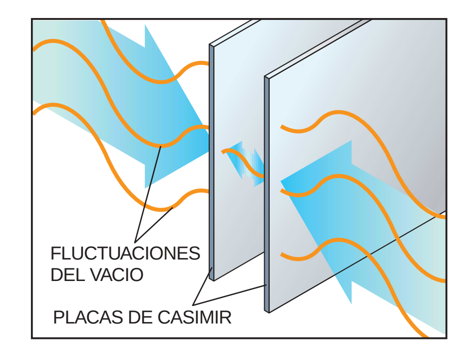
En la teoría general de la relatividad la fuente de las fuerzas gravitatorias (atractivas o repulsivas) es la energía. La materia no es más que una forma de energía. Pero el término cosmológico es distinto. La energía asociada a él no depende de la posición o del tiempo; de ahí que se lo llame “constante cosmológica”. La fuerza que causa actuaría aun cuando no hubiese materia ni radiación. Su fuente, pues, sería una curiosa forma de energía que residiera en el espacio vacío. La constante cosmológica, como el éter, le daba al vacío una especie de aura casi metafísica. Al prescindir de ella la naturaleza volvía a ser razonable.

¿O no? En los años treinta la constante cosmológica dio unas primeras señales de vida en un contexto diferente: el intento de combinar las leyes de la mecánica cuántica con la teoría especial de la relatividad. Paul A. M. Dirac y, luego, Richard Feynman, Julian S. Schwinger y Shinichiro Tomonaga mostraron que el espacio vacío era más complicado de lo supuesto. Resultaba que las partículas elementales podían brotar espontáneamente de la nada y desaparecer de nuevo, siempre y cuando fuese en un tiempo cuya brevedad impidiera la medición. Estas partículas virtuales, así se las conoce, producen efectos mensurables; alteran los niveles de energía de los átomos y crean fuerzas entre las placas metálicas neutras. La teoría de las partículas virtuales concuerda con las observaciones hasta el noveno decimal. Guste o no, el espacio vacío no está, después de todo, vacío.

Si las partículas virtuales pueden cambiar las propiedades de los átomos, ¿podrían también afectar a la expansión del universo? Yakov B. Zeldovich mostró en 1967 que la energía de las partículas virtuales actuaría como la energía asociada a la constante cosmológica. Pero había un serio problema. La teoría cuántica predice todo un espectro de partículas virtuales, que abarcaría todas las longitudes de onda posibles. Al sumar todas las contribuciones sale una energía total infinita. Incluso cuando los teóricos ignoran los efectos cuánticos menores que cierta longitud de onda —más allá de la cual se supone que unos fenómenos gravitatorios cuánticos mal conocidos cambian las cosas—, la energía que le calculan al vacío es unos 120 órdenes de magnitud mayor que la contenida en toda la materia del universo.

¿Qué consecuencias tendría una constante cosmológica tan inflada? Aprovechando la cita de Orwell, es fácil poner un límite observacional a su valor. Mire los dedos de su mano. Si la constante alcanzase el valor que le asigna la teoría cuántica, el espacio entre los ojos y la mano se expandiría tan deprisa que la luz de ésta nunca llegaría a aquéllos. Para ver lo que se tuviese ante la cara habría que luchar sin parar (por así decirlo), pero nunca valdría para nada. El mero hecho de que veamos algo quiere decir que la energía del espacio vacío no puede ser grande. Y que podamos ver, no ya dónde terminan los brazos, sino hasta los remotos confines del universo, impone un límite aún más estricto a la constante cosmológica: ha de ser casi 120 órdenes de magnitud menor que el cálculo mencionado antes. Esta discrepancia entre la teoría y la observación es el problema cuantitativo más desconcertante de la física actual.






**4. DEMOSTRACION experimental del efecto Casimir, medio para corroborar la teoría de que el espacio está lleno de efímeras “partículas virtuales”. El efecto genera unas fuerzas entre los objetos metálicos, de atracción entre placas metálicas paralelas (*arriba*), por ejemplo. Hablando sin mucho rigor, el hueco finito entre las placas impide que se materialicen partículas virtuales con una longitud de onda mayor que una determinada. Por tanto, hay más partículas fuera de las placas que entre ellas, desequilibrio que las empuja entre sí y acerca (*derecha*). El efecto Casimir depende de la morfología de las placas, gracias a lo cual es posible diferenciarlo de las demás fuerzas de la naturaleza.**

La respuesta más a mano es que habría alguna ley física no descubierta que anularía la constante cosmológica. Mas por mucho que plazca a los teóricos que la constante desaparezca, ciertas observaciones astronómicas —de la edad del universo, de la densidad de la materia en su seno y de la naturaleza de las estructuras cósmicas— abogan por lo contrario.

Entre las cuestiones persistentes de la cosmología se incluye la edad del universo. Midiendo la velocidad de las galaxias, podemos calcular cuánto han tardado en ocupar su posición presente, suponiendo que todas partieron del mismo punto. En una primera aproximación podemos prescindir de la deceleración causada por la gravedad: el universo se expande a velocidad constante y el intervalo de tiempo transcurrido es el cociente entre la distancia intergaláctica y la velocidad de separación medida, es decir, la inversa de la constante de Hubble. Cuanto mayor sea ésta, más rápida será la expansión y más joven el universo.

La primera evaluación que Hubble hizo de la constante que lleva su nombre fue de casi 500 kilómetros por segundo por megaparsec; con ello indicaba que dos galaxias separadas por una distancia de un megaparsec (unos tres millones de años luz) se separarían, en promedio, a 500 kilómetros por segundo. Dado este valor, el cosmos tendría unos 2000 millones de años, en penosa contradicción con la edad comprobada de la Tierra, cifrada en unos 4000 millones de años.

LAWRENCE M. KRAUSS, catedrático de la Universidad Case Western Reserve, trabaja en la zona donde convergen física y astronomía. Estudia la naturaleza de estrellas, agujeros negros, lentes gravitatorias y universo primitivo.

Pero si se toma en consideración la atracción gravitatoria de la materia, el análisis predice que los objetos se movieron más deprisa al principio y, por tanto, que tardaron menos en llegar a su posición actual que si la velocidad hubiera sido constante. Esta precisión reduce la edad calculada en un tercio, con lo que la discrepancia resulta desastrosamente peor.

A lo largo de los últimos setenta años se ha determinado mejor la velocidad de la expansión del universo, aunque persistiera la tensión entre la edad que se le calcula y la de algunos de sus objetos. Pero desde hace un tiempo han empezado a converger las diversas mediciones de la constante de Hubble gracias al lanzamiento del *Telescopio Espacial Hubble* y al desarrollo de nuevas técnicas de observación. Wendy L. Freedman, de los Observatorios Carnegie, y sus colaboradores le conceden un valor de 73 kilómetros por segundo por megaparsec (con un intervalo más probable, dependiendo del error experimental, de 65 a 81). Estos resultados ponen el límite superior de la edad de un universo plano en unos 10.000 millones de años.

¿Es suficiente esa edad? Depende de la que tengan los cuerpos más antiguos que daten los astrónomos. Las estrellas más viejas de nuestra galaxia se refugian en los cúmulos globulares, ubicados algunos de éstos en los aledaños de la Vía Láctea, razón por la cual se sospecha que nacieron antes que el resto de la Vía Láctea. Las evaluaciones de su edad, basadas en el cálculo del ritmo con que las estrellas queman su combustible nuclear, oscilan entre los 15.000 y los 20.000 millones de años. Serían, pues, más viejas que el universo.

Para resolver si el conflicto de las edades debíase a fallos de la teoría cosmológica o a los modelos estelares, mis colaboradores Brian C. Chaboyer, Pierre Demarque y Peter J. Kernan y yo reevaluamos en 1995 las edades de los cúmulos globulares. Simulamos los ciclos de vida de tres millones de estrellas cuyas propiedades cubrían las aporías existentes. Comparamos luego nuestras estrellas modelo con las de los cúmulos globulares. Las más viejas, concluimos, tendrían no más de 12,5 millones de años. Pero esa cifra seguía sin concordar con la edad de un universo plano dominado por la materia.

Hasta que, hace dos años, el satélite *Hiparcos*, lanzado por la Agencia Espacial Europea para determinar la ubicación de más de 100.000 estrellas cercanas, revisó las distancias a que se encontraban e, indirectamente, las distancias de los cúmulos globulares. Los nuevos valores alteraron las estimaciones admitidas de su brillo. Nos forzaron a rehacer nuestro análisis, porque el brillo determina la velocidad a la que las estrellas consumen el combustible y, por tanto, su esperanza de vida. Parece ahora que los cúmulos globulares podrían tener, en el límite de los márgenes de error, sólo 10.000 millones de años, lo que ya encaja con las edades cosmológicas.

Ahora bien, esa concordancia marginal no deja de resultar incómoda. Exige que ambos conjuntos de edades estimadas se hallen cerca del borde de los intervalos admisibles. Lo único que puede abandonarse es la suposición de que vivimos en un universo plano dominado por la materia. Una densidad menor de materia, que correspondería a un universo abierto con una deceleración menor, aliviaría un tanto la tensión. Aun así, la única forma de elevar la edad por encima de los 12.500 millones de años sería aceptar que el universo no está dominado por materia, sino por una constante cosmológica. La fuerza repulsiva resultante haría que la expansión de Hubble se acelerara con el tiempo. Las galaxias habrían estado distanciándose antaño menos céleres que hoy y habrían tardado más en ocupar su situación actual, por lo que el universo sería más viejo.

Los cálculos que se manejan sobre la edad del universo son sólo indicativos. Otros pilares de la cosmología observacional se han visto sacudidos también. Al escrutar zonas cada vez mayores del universo ha mejorado la capacidad de tomar nota de su contenido. Contamos con datos convincentes de que la cuantía total de materia que encierra el universo no basta para hacerlo plano.

En ese censo cósmico se calcula primero la síntesis de elementos en la gran explosión originaria. Los elementos ligeros del universo —hidrógeno, helio y sus isótopos menos abundantes, como el deuterio— se crearon en los primeros tiempos del universo en unas cantidades relativas que dependieron del número de protones y neutrones disponibles, los constituyentes de la materia normal. Comparando, pues, la concentración relativa de los distintos isótopos se infiere la cantidad total de materia ordinaria que se produjo en la gran explosión. (Ni que decir tiene que pudo haber habido otra materia cuya composición no fuera de protones y neutrones.)

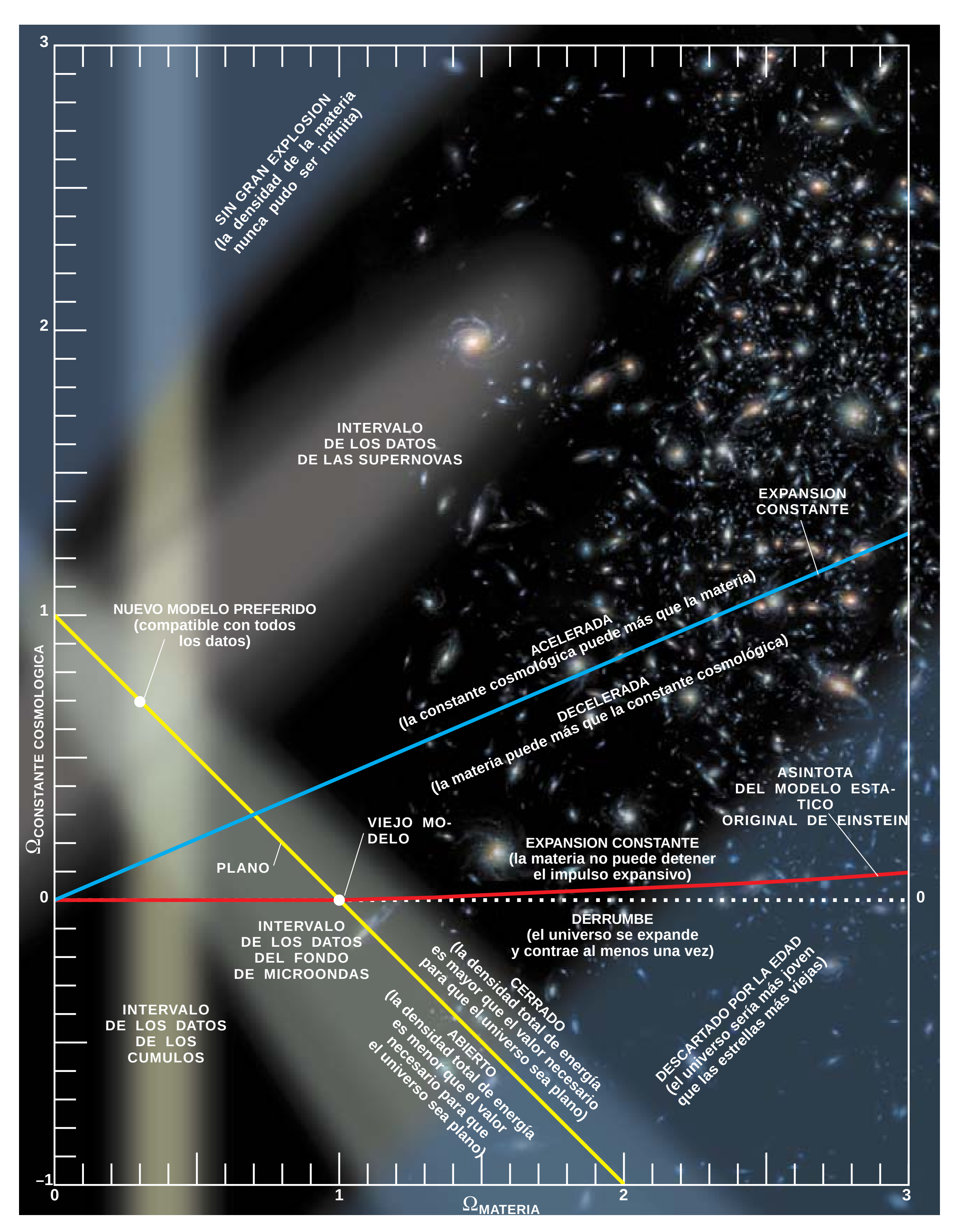
Las observaciones pertinentes dieron un paso de gigante en 1996. El grupo dirigido por David R. Tytler y Scott Burles midieron la abundancia primordial del deuterio observando la absorción de la luz de los cuásares por las nubes de hidrógeno intergalácticas. Al tratarse de nubes que nunca alojaron estrellas, su deuterio debía proceder sólo de la gran explosión. De los descubrimientos de Tytler y Burles se desprende que la densidad media de la materia ordinaria está entre el 4 y el 7% de la necesaria para que el universo sea plano.

Se ha sondeado también la densidad de materia a través del estudio de las principales macroestructuras mantenidas por gravedad, los cúmulos galácticos. Estas agrupaciones de cientos de galaxias abarcan casi toda la materia visible. La mayor parte de su contenido luminoso es un gas intergaláctico caliente que emite rayos X. Su temperatura, inferida del espectro de éstos, depende de la masa total del cúmulo: cuanto mayor sea, mayores son la gravedad y la presión que hace que el gas se aguante contra ella, y esto conduce a una temperatura más alta. El equipo de Simon D. White, tras recoger información acerca de varios cúmulos, llegó a la con-

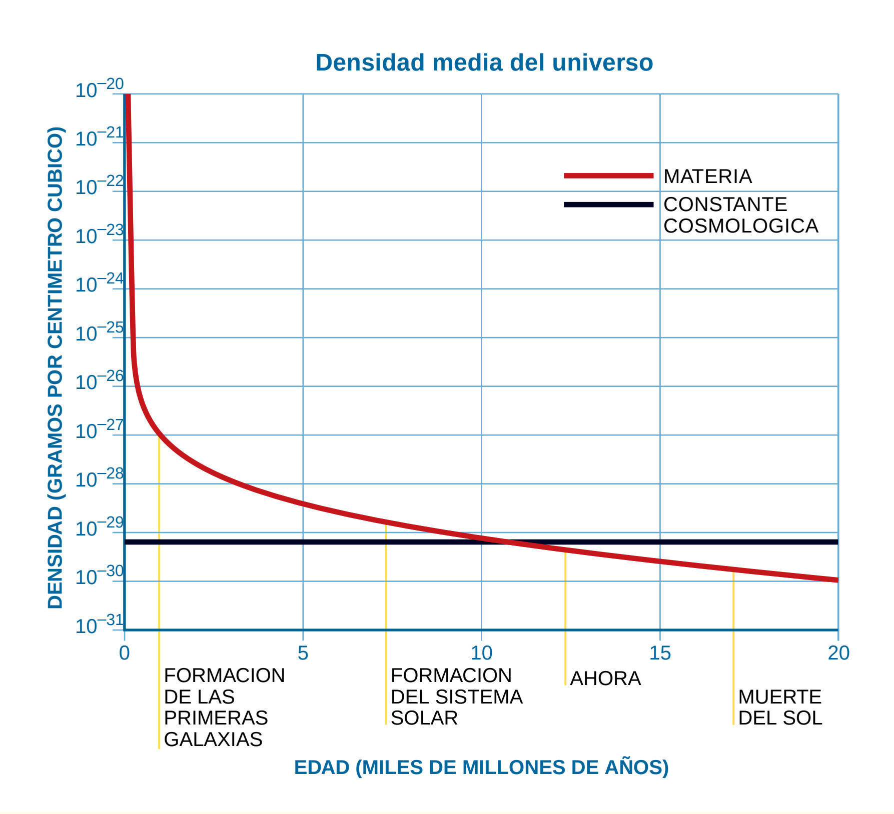
**Sumario de los valores inferidos de la densidad cósmica de materia**

| Observación | $\Omega_{materia}$ |
|---|---|
| Edad del universo | <1 |
| Densidad de protones y neutrones | 0,3–0,6 |
| Cúmulos de galaxias | 0,3–0,5 |
| Evolución galáctica | 0,3–0,5 |
| Radiación del fondo de microondas cósmico | ≤1 |
| Supernova de tipo Ia | 0,2–0,5 |

**5. LAS MEDICIONES de la contribución a Ω de la materia concuerdan más o menos. Aunque cada una de estas mediciones tiene sus escépticos, la mayoría de los astrónomos acepta ahora que la materia sola no puede hacer que Ω sea igual a 1. Pero es posible que otras formas de energía, como la constante cosmológica, aporten lo que falta.**




**6. ESTE MAPA DE LOS MODELOS muestra cómo depende la evolución del universo de dos magnitudes cosmológicas claves: la densidad media de materia (*eje horizontal*) y la densidad de energía en la constante cosmológica (*eje vertical*). Sus valores, dados aquí en unidades cosmológicas estándar, tienen tres efectos distintos. Primero, su suma (que representa el contenido energético cósmico total) determina la geometría del espacio-tiempo (*línea amarilla*). Segundo, su diferencia (que representa las intensidades relativas de la expansión y de la gravedad) determina el cambio de velocidad de la expansión con el tiempo (*línea azul*). Estos dos efectos han sido sondeados por observaciones recientes (*regiones sombreadas*). El tercer factor, un equilibrio de las dos densidades, determina el destino del universo (*línea roja*). Los tres efectos tienen muchas permutaciones, al contrario de lo que sucede si la constante cosmológica es cero.**




**7. LA COINCIDENCIA COSMICA es uno de los muchos misterios que rodean a la constante cosmológica. La densidad media de la materia ordinaria disminuye con la expansión del cosmos (*en rojo*). La densidad equivalente representada por la constante cosmológica es fija (*en negro*). ¿Por qué, pues, las dos tienen hoy casi el mismo valor pese a tales comportamientos antagónicos? Esta concordancia, o es puro azar, una precondición de la existencia humana (lo que remite al principio antrópico débil), o una indicación de que actúa un mecanismo cuya naturaleza no se vislumbra hoy por hoy.**

## El destino del universo

La constante cosmológica cambia la idea más común y sencilla acerca del futuro del universo. La cosmología ha venido prediciendo dos resultados posibles que dependían de la geometría del universo o, equivalentemente, de la densidad media de materia. Si la densidad de un universo lleno de materia sobrepasa cierto nivel crítico, es "cerrado", en cuyo caso acabará por dejar de expandirse, empezará a contraerse y al final desaparecerá en un ardiente apocalipsis. Si es menor que ese valor crítico, será "abierto" y se expandirá eternamente. Un universo "plano", en el que la densidad es igual al valor crítico, se expandirá también para siempre, pero más despacio.

Ahora bien, estos órdenes de cosas parten de que la constante cosmológica sea nula. Si no lo es, puede que sea ella y no la materia la que controle el destino final del universo. La razón estriba en la constante, que, por definición, representa una densidad fija de energía en el espacio. La materia no puede competir: al doblarse el radio su densidad se divide por ocho. En un universo en expansión la densidad de energía asociada a la constante cosmológica lleva las de ganar. Si la constante es positiva, generará en el espacio una fuerza repulsiva de largo alcance y el universo seguirá expandiéndose, aunque la densidad total de la energía en la materia y el espacio supere el valor crítico. (Quedan descartados los valores negativos grandes de la constante porque la resultante fuerza atractiva ya habría puesto fin al universo.)

Esta nueva predicción de la expansión eterna ni siquiera presupone que la constante lo sea en realidad, como reclama la relatividad general. Si la densidad de energía del espacio vacío varía con el tiempo, el destino del universo dependerá de cómo lo haga. Y podría haber un precedente para esos cambios: la expansión inflacionaria del universo primordial. Quizás el universo esté ahora entrando en una nueva era de inflación, que acabaría por llegar a un final. *—L. M. K.*

clusión según la cual la materia luminosa abarcaba entre el 10 y el 20 % de su masa total. De estos resultados, cuando se los combina con las mediciones del deuterio, se desprende que la densidad total de materia incluida en cúmulos —contando protones, neutrones y partículas más exóticas, como ciertas candidatas a materia oscura— no pasa, como mucho, del 60 % de la requerida para que el universo sea plano.

Un tercer conjunto de observaciones, que también tiene que ver con la distribución de la materia a las mayores escalas, respalda la tesis de un universo con masa insuficiente para ser plano. Ningún otro campo de la cosmología ha avanzado tanto en los últimos veinte años como el del conocimiento sobre el origen y la naturaleza de las estructuras cósmicas. Se da por sentado desde hace mucho que las galaxias se agregaron a partir de ligeras concentraciones de materia en el universo primitivo, pero nadie sabía el agente de tales ondulaciones. El desarrollo de la teoría inflacionaria en los años ochenta aportó el primer mecanismo verosímil: el reforzamiento de las fluctuaciones cuánticas hasta que adquirieron un tamaño macroscópico.

Las simulaciones numéricas del crecimiento de las estructuras tras la inflación nos han mostrado que, si la materia oscura, en vez de protones y neutrones, consta de otro tipo de partículas (las llamadas WIMP), entonces ciertas perturbaciones sutiles de la radiación de fondo de microondas pudieron crecer hasta convertirse

en las estructuras que vemos ahora. Además, las concentraciones de materia deberían seguir evolucionando y creando cúmulos de galaxias si hay una densidad global de materia muy alta. El crecimiento parsimonioso del número de cúmulos ricos a lo largo de la historia reciente del universo da a entender que la densidad de materia no llega al 50 por ciento de la necesaria para defender un universo plano.

Este rosario de pruebas de la insuficiencia de materia para pensar en un universo plano vence la arraigada resistencia contra esta posibilidad. Así las cosas, caben dos interpretaciones. Estamos ante un universo abierto, la primera; la segunda, estamos ante un universo plano gracias a alguna forma adicional de energía que no está asociada a la materia ordinaria.

Para optar por una u otra opción se ha pugnado por medir el fondo de microondas con una gran resolución. Las investigaciones provisionales respaldan en estos momentos un universo plano. Mientras, los investigadores que estudian las supernovas remotas han proporcionado la primera prueba directa, aunque aproximada, de la aceleración de la expansión del universo, un signo que apunta a una constante cosmológica con el mismo valor que se desprende de los demás datos.

Las observaciones del fondo de microondas y las de las supernovas iluminan dos aspectos diferentes de la cosmología. El fondo de microondas revela la geometría del universo, que es sensible a la densidad total de energía, sea cual sea su forma, mientras que las supernovas sondean directamente la velocidad de expansión, que depende de la diferencia entre la densidad de materia (que frena la expansión) y la constante cosmológica (que puede acelerarla).

La combinación de todos estos resultados nos da a entender que la constante aporta del 40 al 70 por ciento de la energía necesaria para que el universo sea plano. Los teóricos están ya dándole vueltas a algo que hace 20 años habría sido impensable: una constante cosmológica mayor que cero, pero mucho menor que la predicha por la actual teoría cuántica. Un prodigio de sintonía fina ha de eliminar las energías de las partículas virtuales hasta el lugar decimal 123, pero dejando intacto el 124, una precisión no vista en ninguna otra parte de la naturaleza.

Steven Weinberg y sus colaboradores se ha adentrado en una línea que bascula sobre el último recurso de cosmólogos, el principio antrópico. Si el universo observado es uno entre una infinitud de universos sin vinculación mutua —cada uno de los cuales podría tener constantes de la naturaleza ligeramente distintas, según han venido a postular recientes encarnaciones de la teoría inflacionaria combinada con ideas emergentes de gravedad cuántica—, entonces los físicos pueden confiar en calcular la magnitud de la constante cosmológica inquiriendo en qué universos podría darse vida inteligente. Weinberg y otros han llegado a un resultado que es coherente con la cuantía aparente de la constante cosmológica hoy.

Pero la mayoría de los teóricos no están por esa labor. En su opinión, no hay razón por la que la constante deba tomar un valor determinado; lo toma, simplemente. Aunque este argumento pudiera ser cierto, no se han agotado otras posibilidades, que podrían restringir el valor de la constante mediante una teoría fundamental y no por un accidente histórico.

En otra línea de investigación milita la tradición iniciada por Dirac. Argumentaba éste que se ha medido un número grande del universo, su edad (o, equivalentemente, su tamaño). Si ciertas magnitudes físicas cambiasen con el tiempo hoy podrían ser o muy grandes o muy pequeñas. La constante cosmológica podría ser uno de esos casos. Podría, en efecto, no ser constante. Al fin y al cabo, si la constante cosmológica es fija y no vale cero, es que estamos viviendo en el primer y único momento de la historia cósmica en el que la densidad de materia, que disminuye con la expansión del universo, es equiparable a la energía almacenada en el espacio vacío. ¿Por qué esa coincidencia? Ciertos grupos, en cambio, han imaginado que alguna forma de energía cósmica imita una constante cosmológica, aunque varía con el tiempo.

Se adentraron por esa senda P. James Peebles y Bharat V. Ratra, hace diez años. Movidos por los nuevos hallazgos sobre las supernovas, otros grupos la han resucitado. Algunos se han inspirado en los conceptos de la teoría de cuerdas. Robert Caldwell y Paul J. Steinhardt han vuelto a proponer la denominación “quintaesencia” para nombrar esa energía variable. Da una idea del revuelo teórico de hoy el que la materia oscura que en un principio fue llamada así parezca ahora casi corriente y moliente en comparación. Por mucho que me guste el nombre, ninguna de las ideas concebidas para esa quintaesencia parece convincente. Todas son *ad hoc*. El problema de la constante cosmológica sigue retándonos.

¿Cómo sabrán los cosmólogos que se han reconciliado con este universo que tanta perplejidad teórica causa? Nuevas mediciones del fondo de microondas, el análisis continuo de las supernovas remotas y la medición de las lentes gravitatorias creadas por los cuásares lejanos deberían permitir que se precisase el valor de la constante cosmológica en unos años. Una cosa ya es segura. La cosmología vigente en los años ochenta, según la cual el universo era plano y estaba dominado por materia, ha muerto. El universo o es abierto o está lleno de una energía de origen desconocido. Aunque creo que las observaciones apuntan en favor del segundo estado de cosas, ambos impondrían una visión de la física radicalmente nueva.

BIBLIOGRAFIA COMPLEMENTARIA

DREAMS OF A FINAL THEORY. Steven Weinberg, Pantheon Books, 1992.

PRINCIPLES OF PHYSICAL COSMOLOGY. P. James E. Peebles. Princeton University Press, 1993.

BEFORE THE BEGINNING: OUR UNIVERSE AND OTHERS. Martin Rees. Addison-Wesley, 1997.

THE AGE OF GLOBULAR CLUSTERS IN LIGHT OF HIPPARCOS: RESOLVING THE AGE PROBLEM? Brian Chaboyer, Pierre Demarque, Peter J. Kernan y Lawrence M. Krauss en *Astrophysical Journal*, vol. 494, n.º 1, págs. 96-110, 10 de febrero de 1998. Preprint disponible en xxx.lanl.gov/abs/astro-ph/9706128, de la World Wide Web.

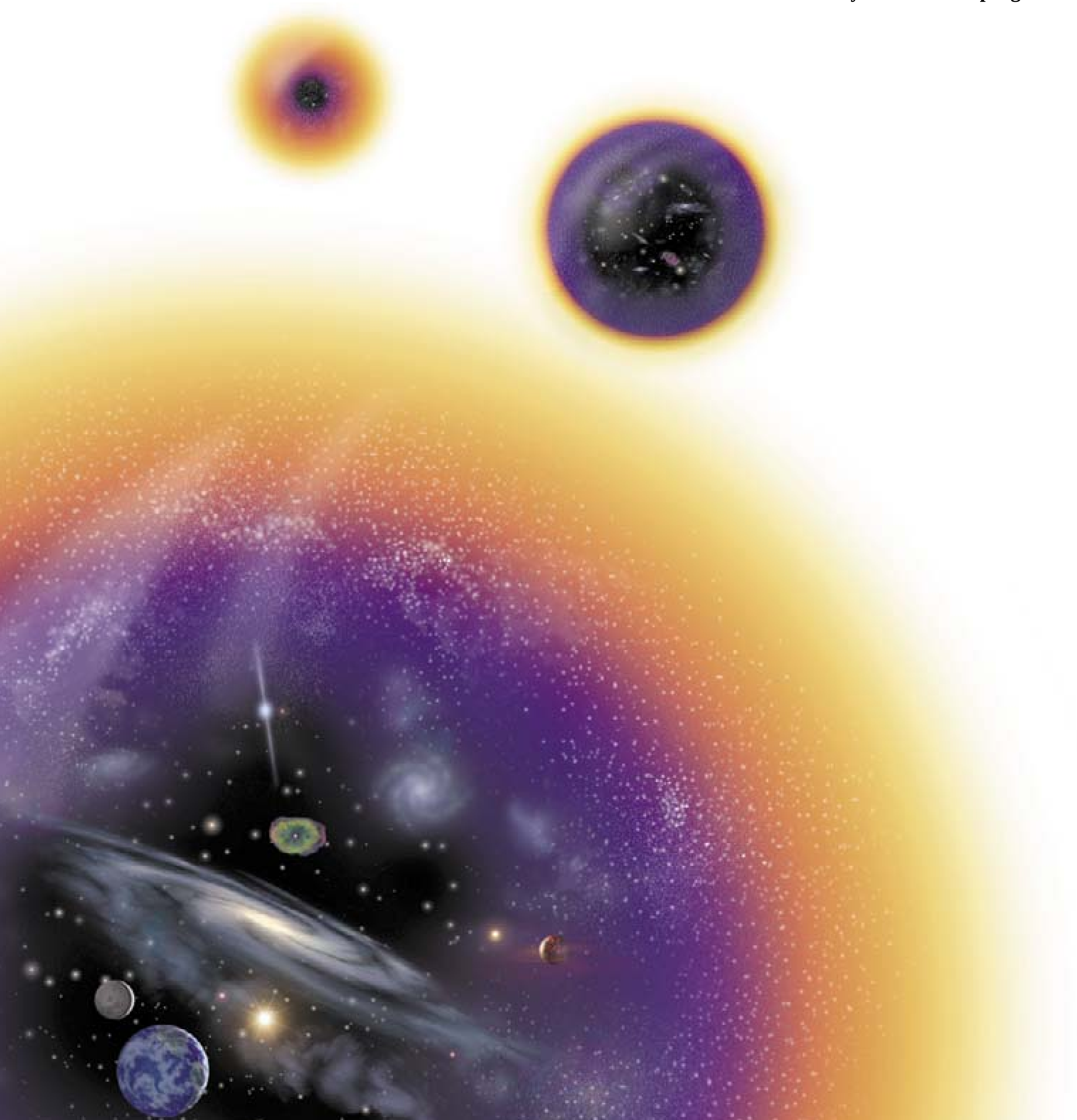
THE END OF THE AGE PROBLEM, AND THE CASE FOR A COSMOLOGICAL CONSTANT REVISITED. Lawrence M. Krauss en *Astrophysical Journal*, vol. 501, n.º 2, págs. 461-466, 10 de julio. Preprint disponible en xxx.lanl.gov/abs/astro-ph/9706227, de la World Wide Web.

LIVING WITH LAMBDA. J. D. Cohn. Preprint disponible en xxx.lanl.gov/abs/astro-ph/9807128, de la World Wide Web.

# Inflación en un universo

*Las pruebas recabadas indican que el universo contiene menos materia y, por tanto, se expande con mayor celeridad, de lo predicho por la teoría inflacionaria. Un refinamiento del modelo, sin embargo, ofrece una explicación coherente de las observaciones*

**Martin A. Bucher y David N. Spergel**



# de baja densidad

La cosmología tiene reputación de ciencia difícil. Pero en muchos aspectos resulta harto más sencillo explicar el universo entero que entender un ser unicelular. A escala cósmica, donde estrellas, galaxias e incluso cúmulos galácticos son meros puntos, la materia se distribuye uniforme. Y está gobernada por la sola fuerza de la gravedad. Sobre esas dos observaciones básicas —la uniformidad a escala cósmica y el predominio de la gravedad— se asienta la teoría de la gran explosión, según la cual nuestro universo se halla inmerso en un proceso de expansión desde hace 12.000 millones de años más o menos. Pese a tan sencillos pilares, la teoría explica con notable coherencia la velocidad de recesión mutua entre galaxias, la concentración relativa de elementos ligeros, el tenue brillo de microondas del firmamento y la evolución general de las estructuras astronómicas. Diríase que la evolución del cosmos ni se inmuta con los avatares de sus componentes. Para desazón de biólogos, no podemos aplicar el mismo principio ni al más simple de los organismos.

Con todo, la teoría de la gran explosión no está exenta de paradojas. Veinte años atrás los cosmólogos resolvieron incómodas contradicciones incorporando ideas tomadas de la física de partículas. Así nació la teoría de la "inflación". Ahora, sin embargo, es ésta la que debe hacer frente a una crisis, provocada por observaciones recientes que contradicen su predicción en torno a la densidad media de materia en el cosmos. Se percibe que el universo quizá no sea tan simple como se le había imaginado. Debe elegirse entre admitir la existencia de una forma exótica de materia o energía, o bien añadir otro nivel de complejidad a la teoría de la inflación. En el artículo nos concentraremos en la segunda opción.

Hablando con propiedad, la teoría de la gran explosión no describe el nacimiento del universo, sino su desarrollo hasta adquirir forma de estado adulto. En ese contexto, el universo bebé era una suerte de densa caldera de radiación extremadamente caliente. Una fracción de la misma, no mayor que el tamaño de un nabo tierno, se desarrolló hasta convertirse en el universo observable hoy. (Hay otras partes del universo, de infinita extensión tal vez, que se nos escapan porque su luz no ha tenido tiempo todavía de llegar hasta la Tierra.) La idea de un universo en expansión puede resultar extraña. El propio Albert Einstein la recibió receloso en un principio. Cuando el cosmos se expande, aumenta la distancia entre dos objetos cualesquiera. Las galaxias remotas se alejan porque el espacio interpuesto crece.

**1. LOS UNIVERSOS DE BURBUJA, autocontenidos, crecen dentro de un "multiverso" mayor y vacío. En coherencia con los postulados fascinantes que la relatividad insta, tiempo y espacio adquieren diferentes significados dentro y fuera de cada burbuja. Considerado desde el interior, el tiempo aumenta hacia el centro de la burbuja; la pared de la burbuja representa la gran explosión de dicho universo. Ni que decir tiene que semejante cuadro pergeña una perspectiva imposible. Aun cuando un observador pudiese existir fuera de la burbuja, no podrían mirar hacia dentro, porque la burbuja se expande a la velocidad de la luz.**

Consecuencia natural de la expansión de un universo uniforme es la ley de Hubble. Nos dice que las galaxias se alejan de la Tierra (o de cualquier otro punto de referencia) con velocidades proporcionales a su distancia. No todos los objetos del universo obedecen dicha ley, porque la atracción gravitatoria se opone a la expansión; por ejemplo, el Sol y la Tierra no se alejan. Sí se cumple a escala cósmica. En el marco del modelo más simple de la gran explosión, la expansión siempre ha procedido a la misma velocidad, más o menos.

A medida que el joven universo se expandía, también se enfriaba, se diluía y se volvía cada vez más complejo. Parte de la radiación se condensó en partículas elementales y núcleos atómicos. En unos 300.000 años, la temperatura había bajado hasta los 3000 grados Celsius, que facilitaba la formación de átomos de hidrógeno por combinación de electrones y protones. En ese momento el universo se tornó transparente, liberando radiación cósmica de fondo de microondas. Esa radiación es muy uniforme, prueba de que entonces la densidad de materia en diferentes regiones del universo primitivo variaba sólo de una parte en 100.000. Por sutiles que fuesen esas diferencias, las leves concentraciones se fueron agrandando hasta constituir galaxias y cúmulos galácticos.

Pese a su potencia explicativa, la teoría estándar de la gran explosión no puede dar razón de algunas cuestiones fundamentales. ¿Por qué es tan uniforme el universo? Dos regiones de polos opuestos del firmamento parecen básicamente iguales, aunque disten más de 24.000 millones de años luz. La luz lleva viajando unos 12.000 millones de años; por tanto, esas regiones no se han visto mutuamente todavía. No ha habido tiempo suficiente para que materia, calor o luz fluyesen entre una y otra e igualaran su respectiva densidad y temperatura. En consecuencia, la uniformidad del universo hubo de preceder a la expansión. La teoría estándar, sin embargo, no sabe cómo.

Planteado a la inversa: ¿por qué presentaba el universo primitivo pequeñas variaciones de densidad? Y fue una suerte que las tuviera. Sin esas pequeñas ondulaciones, el universo seguiría con una densidad uniforme —de varios átomos por centímetro cúbico— y ni la Vía Láctea ni la Tierra existirían.

Finalmente, ¿por qué la velocidad de expansión es la que es, ni más ni menos que la suficiente para contrarrestar el efecto de atracción ejercido por la gravedad conjunta de toda la materia del cosmos? Cualquier desviación significativa de ese equilibrio perfecto se habría intensificado a lo largo del tiempo. Si el ritmo de expansión hubiese ido demasiado rápido, hoy el universo carecería casi por completo de materia. Si la gravedad hubiese sido demasiado intensa, el universo habría terminado ya su existencia en una gran implosión ("big crunch").

Para expresar esa cuestión los cosmólogos se sirven de la variable omega, $\Omega$, el cociente entre la energía gravitatoria y la energía cinética (es decir, la energía contenida en el movimiento de materia conforme se expande el espacio). Esa variable es proporcional a la

MARTIN A. BUCHER y DAVID N. SPERGEL estudian la física del universo primitivo. Bucher, adscrito al departamento de matemática aplicada y física teórica de la Universidad de Cambridge, es uno de los creadores de la teoría inflacionaria abierta. Spergel, profesor del departamento de astrofísica de Princeton, participa en la preparación de la *Sonda de Anisotropías de Microondas*.

densidad de materia del universo —una densidad mayor significa una fuerza de gravedad más intensa y, por tanto, una Ω mayor. Si $\Omega = 1$, su valor nunca cambia. En cualquier otro caso, su valor crecerá o menguará rápidamente en un proceso autoalimentado, a medida que la energía cinética o la gravitatoria empiecen a dominar. Tras miles de millones de años, Ω debería ser muy próxima a cero o a infinito. Como la densidad presente del universo no es —para nuestra fortuna— ni cero ni infinito, el valor original de Ω ha debido ser exactamente uno o casi (dentro de una parte en $10^{18}$). ¿Por qué? La teoría de la gran explosión lo ignora. A no ser que tomemos por explicación la pura chamba.

Tales limitaciones no invalidan el modelo —que da cuenta de miles de millones de años de historia cósmica—, pero denuncian su carácter incompleto. Para llenar esa laguna Alan H. Guth, Katsuhiko Sato, Andrei D. Linde, Andreas Albrecht y Paul J. Steinhardt desarrollaron a principios de los ochenta la teoría de la inflación.

Mas para resolver esas paradojas resulta obligado complicar la teoría de la gran explosión. El modelo inflacionario postula que, en su infancia, el universo atravesó un período de expansión muy rápida ("inflacionaria"). Si la expansión tradicional de la teoría estándar deceleraba con el tiempo, la expansión inflacionaria se aceleraba. Inducía el distanciamiento, cada vez más rápido, de unos objetos de otros. Pese a que la separación llegó a proceder más veloz que la luz, no violaba la relatividad, que prohíbe que los cuerpos de masa finita se muevan por el espacio más deprisa que la luz. Los objetos, en realidad, permanecían quietos en relación con el espacio circundante. Era el propio espacio el que se expandía más rápido que la luz.

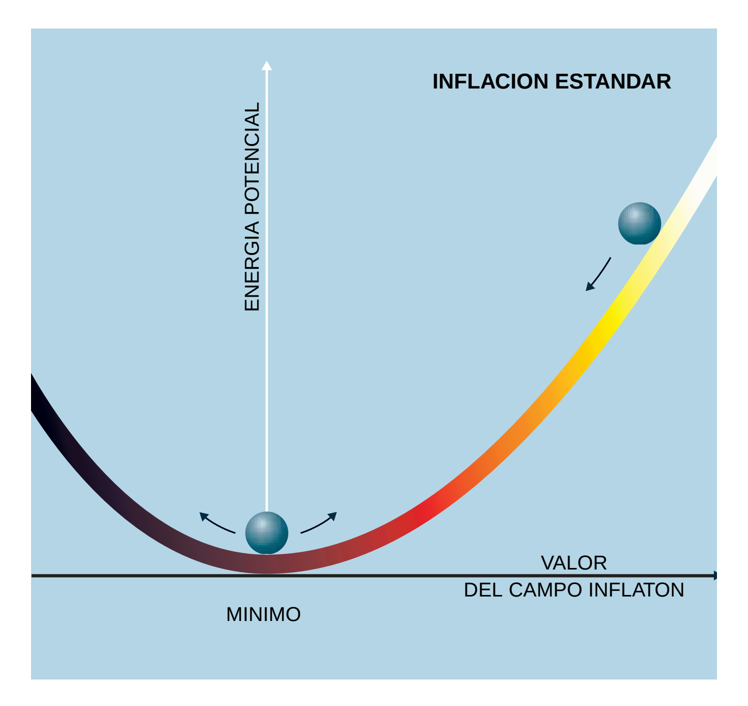
A esa célere expansión primordial hay que atribuir la uniformidad del universo observable hoy, cuyas partes visibles estuvieron tan cerca unas de otras, que pudieron tener una densidad y temperatura comunes. Durante la inflación, diversas partes de ese cosmos uniforme perdieron el contacto; sólo más tarde, concluido el período de inflación, la luz tuvo tiempo de alcanzar el ritmo de la expansión, más lento, que asignamos a la gran explosión. Si persiste alguna inhomogeneidad en el universo, no ha caído todavía en nuestro campo de visión.

Para desencadenar una expansión rápida, la teoría inflacionaria añade un nuevo concepto, extraído de la física de partículas: el campo "inflatón". Protones, electrones y otras partículas elementales están representados por campos cuánticos, similares a los campos eléctrico, magnético y gravitatorio. Un campo no es más que una función del espacio y del tiempo cuyas oscilaciones se describen mediante partículas. Los campos son responsables de la transmisión de las fuerzas.

El campo inflatón imparte una fuerza antigravitatoria que estira el espacio. Asociada con un valor dado del inflatón, existe una energía potencial. A la manera de una bola que rueda por una pendiente, el campo inflatón tiende a rodar hacia el mínimo de su potencial. Pero la expansión del universo introduce una suerte de fricción cosmológica, que frena el descenso. Mientras domina ese rozamiento, el inflatón permanece clavado en su sitio. Su valor es casi constante y, por tanto, la fuerza antigravitatoria va ganando con respecto a la gravedad, provocando que la distancia entre objetos inicialmente vecinos aumente a una velocidad creciente. Con el transcurso del tiempo el campo se debilita y convierte el resto de su energía en radiación. A continuación, la expansión del universo sigue la pauta que le marca la teoría estándar de la gran explosión.

Para la representación plástica del proceso se acude a la morfología del cosmos. De acuerdo con la teoría general de la relatividad, la gravedad ejerce un efecto geométrico. Materia y energía curvan el tejido del espacio y el tiempo, distorsionando los caminos que los objetos siguen. La expansión conjunta del universo, que es en sí misma un tipo de curvatura del espacio y del tiempo, está controlada por el valor de Ω. Si Ω es mayor que uno, el universo tiene curvatura positiva, la de la superficie de una naranja aunque en tres dimensiones espaciales (geometría esférica, o cerrada). Si Ω es menor que uno, el universo tiene curvatura negativa, la de una patata frita (geometría hiperbólica, o abierta). Si es igual a uno, el universo es plano, como una torta (geometría euclídea usual).

La inflación aplana el universo observable. Cualquiera que sea la forma inicial del cosmos, la rápida expansión lo infla hasta tamaño colosal y empuja la mayor parte del mismo fuera de nuestra vista. La fracción visible puede parecer plana, igual que una extensión pequeña de la superficie de la Tierra lo parece también. Así pues,



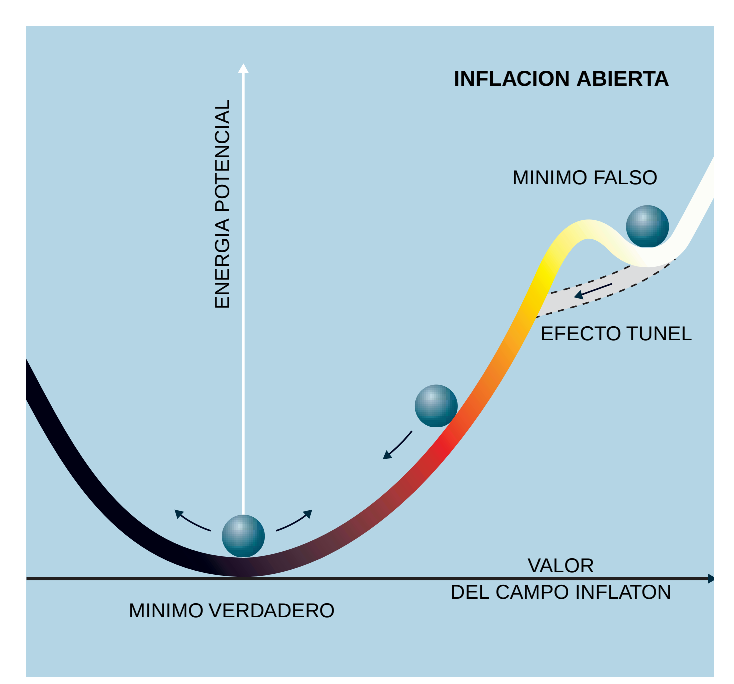

**2. El CAMPO INFLATON, responsable de la fuerza que causó la expansión del espacio, se comporta cual una bola que rodara pendiente abajo: minimiza su energía potencial (*eje vertical*) cambiando su valor (*eje horizontal*). El campo empezó en lo alto de la pendiente, creado por procesos cuánticos en los albores del tiempo. En el marco de la inflación estándar (*izquierda*), el campo rodó directamente hasta su valor mínimo. Pero en el marco de la inflación abierta (*derecha*), quedó atrapado en un valle, o "vacío**

la inflación empuja el valor observado de Ω hacia uno. Al mismo tiempo, cualquier irregularidad inicial en la densidad de materia y radiación se suaviza.

En la teoría inflacionaria estándar, la planitud cósmica y la uniformidad se hallan vinculadas. Para que el universo, afirma el modelo, adquiera la homogeneidad que posee, debe ser planísimo, con Ω igual a uno con precisión de una parte en 100.000. Cualquier desviación de esa lisura absoluta tendría que resultar imperceptible para los astrónomos. A lo largo de los últimos 20 años, se ha venido aceptando que la planitud observacional constituía un postulado firme de la teoría.

Y ahí anida el problema. Toda una gavilla de observaciones, relativas a cúmulos galácticos y supernovas distantes, sugieren ahora que la gravedad es demasiado débil para combatir la expansión. De ser ello cierto, la densidad de materia tendrá que ser menor que la predicha, con Ω igual a 0,3 aproximadamente. Con otras palabras, el universo puede ser curvo y abierto. Hay tres maneras de interpretar ese resultado. A tenor de la primera, la teoría inflacionaria yerra. Pero si hemos de abandonar el concepto de inflación, reaparecerían las paradojas tan elegantemente resueltas por la teoría, y habría que buscar otro modelo. No se conoce alternativa alguna.

Una segunda interpretación ha tomado aliento con la expansión acelerada que se infiere de las observaciones de supernovas remotas. Dicha expansión sugiere la existencia de energía adicional en forma de "constante cosmológica". Esa energía extra actuaría como un extraño tipo de materia, curvando el espacio igual que lo hace la materia ordinaria. El efecto combinado resultante aplanaría el espacio, en cuyo caso la teoría inflacionaria no tiene nada que temer. Pero la inferencia de la constante cosmológica está plagada de interrogantes sobre el polvo cósmico y la naturaleza de las estrellas que explotan en supernova. Las opciones continúan *abiertas* (por traer el agua a mi molino).




**falso". En la mayor parte del universo, se quedó atrapado allí, y la inflación nunca terminó. Pero en algunas el campo escapó de su valle por efecto túnel y completó su descenso. Una de esas regiones se convirtió en la burbuja en la que vivimos. Para los dos modelos inflacionarios, sin embargo, una vez que el campo se acercó a su estadio final de reposo, osciló hacia adelante y hacia atrás, llenando el espacio con materia y radiación. La gran explosion había empezado.**

La tercera interpretación toma las observaciones al pie de la letra. Se pregunta si el universo plano es consecuencia real e inevitable de la inflación. Tal planteamiento reclama una nueva extensión de la teoría a tiempos anteriores incluso, con alguna complejidad añadida. El camino lo señalaron en un comienzo Sidney R. Coleman, Frank de Luccia y J. Richard Gott III, a principios de los ochenta. Ignoradas durante más de una década, esas ideas fueron desarrolladas por uno de nosotros (Bucher), Neil G. Turok, Alfred S. Goldhaber, Misao Sasaki, Takahiro Tanaka y Kazuhiro Yamamoto. Linde y sus colaboradores han propuesto modelos concretos y extensiones de esas ideas.

Si el campo inflatón tuviese una función de energía potencial diferente, la inflación habría curvado el espacio de una manera precisa y predecible; hubiera dejado algo curvado y no plano del todo al universo. Supongamos que la función de energía potencial tuviese dos valles: un mínimo falso (local) y un mínimo verdadero (global). Conforme el inflatón rodó hacia abajo, el universo se expandió y se volvió uniforme. Pero entonces el campo se vio atrapado en el mínimo falso. A ese estado se le denomina "vacío falso"; allí la materia y radiación del cosmos se sustituyeron, casi del todo, por energía del campo inflatón. Las fluctuaciones inherentes a la mecánica cuántica provocaron que el inflatón se tambalease y terminaron por dejarle escapar del mínimo falso, igual que liberamos una bola encallada sacudiendo la máquina del millón.

Esa huida, o desintegración del vacío falso, no ocurrió en todas partes al mismo tiempo. Antes bien, sucedió en algún sitio y luego se propagó. El proceso recordaba el del calentamiento del agua hasta el punto de ebullición, que no se convierte instantáneamente en vapor en todas partes. Debido al movimiento aleatorio de los átomos, primero las burbujas distribuidas al azar forman núcleos por el líquido. Las burbujas inferiores a cierto tamaño mínimo se desploman en virtud de la tensión superficial. Con las burbujas mayores, sin embargo, la diferencia de energía entre el vapor y el agua sobrecalentada supera la tensión superficial; esas burbujas se expanden a la velocidad del sonido en el agua.

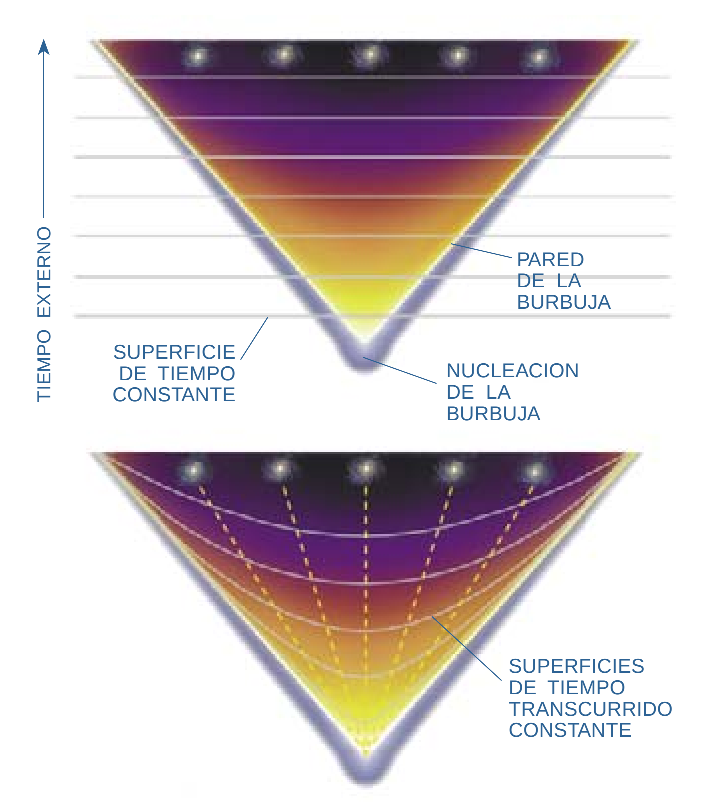
En la desintegración del vacío falso, las fluctuaciones cuánticas desempeñaron la función del movimiento aleatorio de los átomos; provocaron la nucleación de burbujas del vacío verdadero. La tensión superficial destruyó la mayoría de las burbujas, pero algunas consiguieron crecer tanto, que los efectos cuánticos perdieron interés. Sin nada que se opusiera a ello, su radio siguió aumentando a la velocidad de la luz. Conforme la pared externa de la burbuja pasaba por un punto del espacio, el campo inflatón en él saltaba fuera del mínimo falso y retomaba su descenso. A partir de ahí, el espacio interior a la burbuja se hinchó en la celeridad demandada por la teoría inflacionaria estándar. El interior de esa burbuja corresponde a nuestro universo. El momento en que el campo inflatón se desprendió de su mínimo falso corresponde a la gran explosión en teorías precedentes.

Por lo que respecta a puntos situados a diferentes distancias del centro de nucleación, la gran explosión ocurrió en tiempos distintos. Tamaña disparidad resulta dura de aceptar a la primera. Pero un análisis pormenorizado del

campo inflatón revela qué pasó. El inflatón servía de cronómetro: su valor en un punto representaba el tiempo transcurrido desde el instante de la gran explosión en dicho punto. Debido al retraso en el comienzo de la gran explosión, el valor del inflatón no era el mismo en todas partes; era mayor en la pared de la burbuja y caía incesantemente hacia el centro. En un planteamiento matemático, el valor del inflatón era constante sobre superficies en forma de hipérbolas.

El valor del inflatón no es mero constructo teórico. Determina las propiedades básicas del universo encerrado en la burbuja, o sea, su densidad media y la temperatura del fondo de radiación cósmica (hoy 2,7 ºC sobre el cero absoluto). A lo largo de una superficie hiperbólica, la densidad, temperatura y tiempo transcurrido son constantes. Esas superficies son lo que los observadores incluidos en la burbuja perciben como superficies de "tiempo" constante, que difiere del tiempo experimentado fuera de la burbuja.

¿Por qué diverge, de fuera adentro, algo tan fundamental como el tiempo? En una interpretación del espacio y tiempo prerrelativista, tal disparidad resulta imposible. Pero en la teoría einsteiniana, la distinción entre espacio y tiempo se difumina. Lo que cualquier observador denomina "espacio" y "tiempo" es, en buena medida, cuestión de conveniencia. Expresado sin mayor precisión, el tiempo representa la dirección en la cual las cosas cambian; dentro de la burbuja, el cambio está impulsado por el inflatón.




**3. ¿UNIVERSO INFINITO EN UN ESPACIO FINITO? Semejante configuración paradójica es posible. ¿Cómo? El espacio y el tiempo se perciben de manera diferente fuera (*arriba*) y dentro (*abajo*) del universo de burbuja. Aquí, el tiempo —visto por observadores exteriores— transcurre hacia arriba. El espacio, por definición, es cualquier línea o superficie que conecta puntos en determinado tiempo (*líneas horizontales*). La burbuja parece finita. No obstante, los observadores interiores solamente se percatan del tiempo transcurrido, la cantidad de tiempo que ha pasado desde que la pared de la burbuja llegó por primera vez a una posición dada. Conforme el tiempo transcurrido se incrementa, la temperatura decrece, lo que impulsa cambios físicos (*caliente es amarillo; frío es negro*). Las superficies de tiempo transcurrido constante son hipérbolas, que se curvan hacia arriba y nunca alcanzan la pared de la burbuja. Los puntos del interior de la ampolla se alejan entre sí debido a la expansión cosmológica (*líneas de puntos*).**

La relatividad concede al universo tres dimensiones para el espacio y una para el tiempo. Una vez determinada la dirección del tiempo, las tres direcciones restantes serán espaciales; en éstas el tiempo es constante. Por tanto, un universo de burbuja parece hiperbólico desde dentro. Para nosotros, viajar por el espacio exterior no es más que movernos a lo largo de una hipérbola. Mirar hacia atrás en el tiempo es volver la vista hacia la pared de la burbuja. Aunque, en principio, podríamos mirar fuera de la burbuja y antes de la gran explosión, en la práctica el universo primitivo, denso y opaco, obstruye la visión.

Esa mescolanza de espacio y tiempo permite que un universo hiperbólico entero (cuyo volumen es infinito) quepa dentro de una burbuja en expansión (cuyo volumen, aunque ilimitadamente creciente, es siempre finito). El espacio dentro de la burbuja se percibe como una mezcla de espacio y tiempo desde fuera de la burbuja. Lo mismo que el tiempo externo es infinito, también lo es el espacio interno.

La idea de universo burbuja, por extraño que parezca, libera a la teoría inflacionaria de tener que reclamar que Ω sea igual a uno. Aunque la formación de la burbuja creó hipérbolas, nada dijo de su escala precisa. La escala viene determinada por el potencial del inflatón; varía con el tiempo de acuerdo con el valor de Ω. Inicialmente Ω dentro de la burbuja es igual a cero. Durante la inflación su valor se incrementa, acercándose a uno. Las hipérbolas comienzan, pues, con una brusca curvatura para aplanarse poco a poco. El potencial del inflatón determina el ritmo y la duración de ese aplanamiento. Con el tiempo, la inflación que se desarrollaba en el interior de la burbuja termina; ocurre tal cuando Ω está en el umbral de uno, pero sin pisarlo. Luego Ω empieza a decrecer. Si la duración de la inflación en el interior de la burbuja es la correcta (dentro de unos pocos por ciento), el valor presente de Ω coincidirá con el valor observado.

A primera vista, da la impresión de que hemos dibujado un proceso muy retorcido. Atiéndase, sin embargo, a la conclusión principal, que es simple: la uniformidad del universo no tiene por qué guardar relación con la geometría. Antes bien, podrían derivarse de dos estadios diferentes de inflación: la uniformidad, de la inflación antes de la nucleación de la burbuja; la geometría, de la inflación en el interior de la burbuja. Al no estar asociadas esas dos propiedades, la necesidad de uniformidad no determina la duración de la inflación, que procede hasta conferir a las hipérbolas el grado de planitud deseado.

De hecho, hablamos de una extensión directa de la teoría de la gran explosión. La exposición estándar de la inflación describe qué sucedió justo antes de la expansión de la gran explosión. La nueva concepción, conocida como teoría inflacionaria abierta, añade otro estadio que precede a la inflación estándar. Para explicar la creación original del universo se necesitará otra teoría que describa tiempos todavía anteriores.

## La geometría del universo

Si el universo tuviera un "exterior" y pudiésemos asomarnos desde allí a su "interior", la cosmología andaría más segura. Privados de esa atalaya, no queda otro remedio que inferir la forma del universo a partir de sus propiedades geométricas. La experiencia diaria nos dice que el espacio es euclídeo, "plano", a escalas pequeñas: las paralelas no se encuentran nunca, los triángulos recorren 180 grados, la circunferencia de un círculo es $2\pi r$, etc. Pero sería incorrecto suponer que el universo es euclídeo a escalas grandes, igual que lo sería concluir que la Tierra es plana por un segmento de la misma.

Hay otras dos geometrías tridimensionales coherentes con las observaciones de la homogeneidad cósmica (la equivalencia de todos los puntos del espacio) y la isotropía (la equivalencia de todas las direcciones). Son la geometría esférica, o "cerrada", y la geometría hiperbólica, o "abierta". Las dos están caracterizadas por una longitud de curvatura análoga al radio de la Tierra. Si la curvatura es positiva, la geometría es esférica; si negativa, hiperbólica. Para distancias mucho menores, todas las geometrías parecen euclídeas.

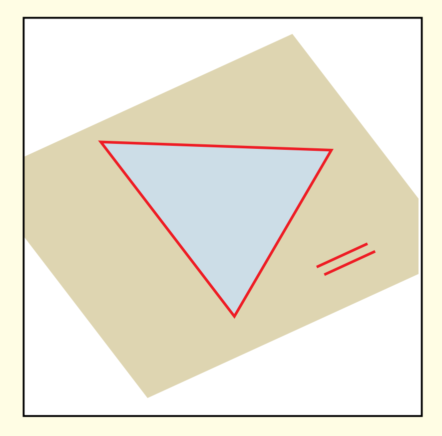
En un universo esférico, las líneas paralelas al final se encuentran, los triángulos recorren hasta 540 grados y la circunferencia de un círculo es menor que $2\pi r$. Puesto que el espacio se repliega sobre sí mismo, el universo esférico es finito. En un universo hiperbólico, las paralelas divergen, los triángulos miden menos de 180 grados y la circunferencia de un círculo es mayor que $2\pi r$. Tal universo, al igual que el espacio euclídeo, es infinito en tamaño.

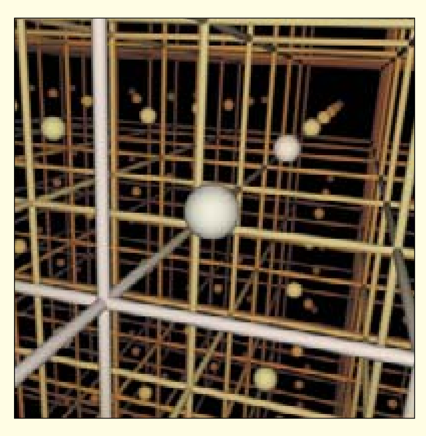
Esas tres geometrías tienen efectos distintos sobre la perspectiva (*derecha*), que distorsiona la apariencia de las características del fondo de radiación de microondas. Las mayores ondulaciones en el fondo tienen el mismo tamaño absoluto, con independencia del proceso de inflación específico. Si el universo es plano, las ondulaciones mayores parecerían tener el tamaño de un grado. Pero si el universo es hiperbólico, las mismas características parecerían alcanzar la mitad de este tamaño, en razón de la distorsión geométrica de los rayos de luz.

Según unas primeras observaciones, las ondulaciones efectivamente miden un grado. Si se confirman, los resultados rechazarían la teoría inflacionaria abierta. Pero no podremos decidir nada hasta que las observaciones desde satélites nos ofrezcan una respuesta definitiva.

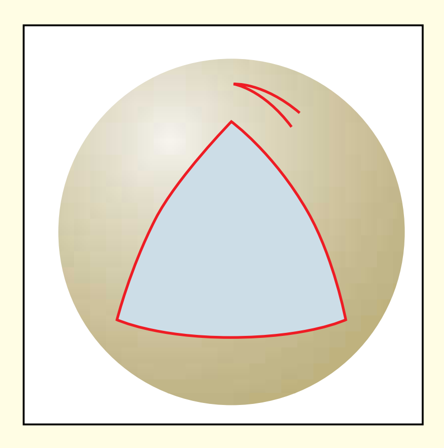
*—M.A.B. y D.N.S.*

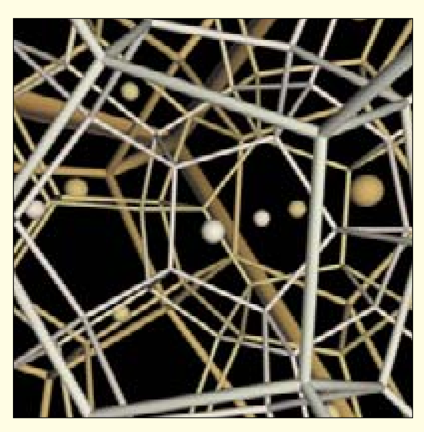
**TRES GEOMETRIAS ofrecidas desde dos perspectivas diferentes: una hipotética visión exterior que ignora, con fines ilustrativos, una de las dimensiones espaciales (*columna de la izquierda*) y una visión interior que muestra las tres dimensiones así como un sistema de referencia (*columna de la derecha*). La visión exterior es útil para representar las reglas geométricas básicas. La visión interior revela los tamaños aparentes de los objetos (que en esos diagramas tienen el mismo tamaño intrínseco) a diferentes distancias. Aquí los objetos y sistemas de referencia enrojecen con la distancia.**



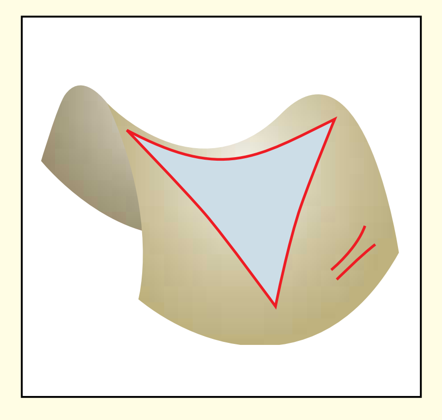


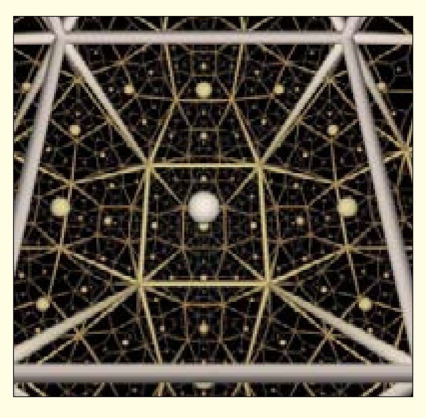
**El espacio plano cumple las reglas de la geometría euclídea. El tamaño angular de esferas idénticas es inversamente proporcional a la distancia: la perspectiva usual del punto que se desvanece, que se enseña en clases de dibujo.**





**El espacio esférico tiene las propiedades geométricas de un globo. Con distancia creciente, las esferas al principio parecen pequeñas. Alcanzan un tamaño aparente mínimo y después parecen mayores. (Del mismo modo, las líneas de longitud que emanan de un polo se separan, llegan a un alejamiento máximo en el ecuador y luego se refocalizan en el polo opuesto.) Este sistema de referencia consiste en dodecaedros.**





**El espacio hiperbólico tiene la geometría de una silla de montar. El tamaño angular disminuye mucho más rápido con la distancia que en el espacio euclídeo. Por ser más agudos los ángulos, cinco objetos cúbicos encajan alrededor de cada vértice y no cuatro sólo.**

## Así empezó el universo

Las leyes de la física describen la evolución seguida por un sistema a partir de un estado inicial. Pero cualquier teoría que explique el origen del universo requiere una pauta distinta, se le exige que dé cuenta del propio estado inicial. Si las leyes normales vienen a ser mapas de carreteras que nos señalan la trayectoria de A a B, las nuevas leyes deben justificar por qué empezamos precisamente en A.

En 1983 James B. Hartle y Stephen W. Hawking aplicaron la mecánica cuántica al universo. Presentaron una función de ondas semejante a la función de ondas para átomos y partículas elementales. La función de ondas determina las condiciones iniciales del universo. Desde ese enfoque, la distinción entre futuro y pasado se desmorona en el universo primitivo; la dirección temporal toma propiedades de una dirección espacial. Del mismo modo que no hay una frontera del espacio, tampoco hay un principio del tiempo identificable como tal. En una hipótesis alternativa, Alexander Vilenkin propuso una función de onda de "efecto túnel", determinada por las probabilidades relativas de que un universo de tamaño cero se convirtiese por sí solo en un universo de tamaño finito.

Hawking y Neil G. Turok sugirieron la creación espontánea de un universo burbuja a partir de la nada. Esa nueva versión de la inflación abierta evita la necesidad de la desintegración de un vacío falso, pero Vilenkin y Andrei D. Linde han cuestionado las premisas introducidas en el cálculo.

Linde ha intentado esquivar el problema de condiciones iniciales: la inflación sería un proceso sin inicio. En la imagen clásica, la inflación termina conforme el inflatón cae rodando potencial abajo. Ahora bien, debido a fluctuaciones cuánticas, el campo puede saltar potencial arriba, además de saltar potencial abajo. De ese modo, habría siempre regiones del universo —la mayor parte del volumen del mismo— que se están inflando. Esas regiones rodean bolsas de espacio donde la inflación ha terminado y un universo estable se ha desarrollado. Cada bolsa tiene un conjunto diferente de constantes físicas; nosotros vivimos en aquella cuyas constantes son adecuadas para nuestra existencia. El resto del universo sigue inflándose y siempre lo ha hecho. Pero Vilenkin y Arvind Borde han sostenido que incluso esa extensión de la inflación no describe completamente el origen del universo. Aunque la inflación puede ser eterna hacia el futuro, en el fondo requiere un comienzo.

J. Richard Gott III y Li-Xin Li defienden que el universo se halla atrapado en un estado cíclico, como una viajera en el tiempo que vuelve hacia atrás y se convierte en su propia madre. Esa persona, sin árbol genealógico ahora, jamás podrá reclamar su origen. En la hipótesis de Gott y Li, nuestra burbuja se desprendió del proto-universo cíclico; ya no es cíclico, sino que siempre se expande y se enfría.

Resulta muy difícil (aunque no imposible quizá) que los astrónomos sometan a prueba cualquiera de estas hipótesis. La inflación borra casi todas las huellas observacionales de lo que la precedió. Muchos sospechan que una explicación del universo preinflacionario —y del origen de las mismas leyes físicas— tendrá que esperar a una teoría fundamental del la física, tal vez la teoría de cuerdas.

*—M.A.B. y D.N.S.*

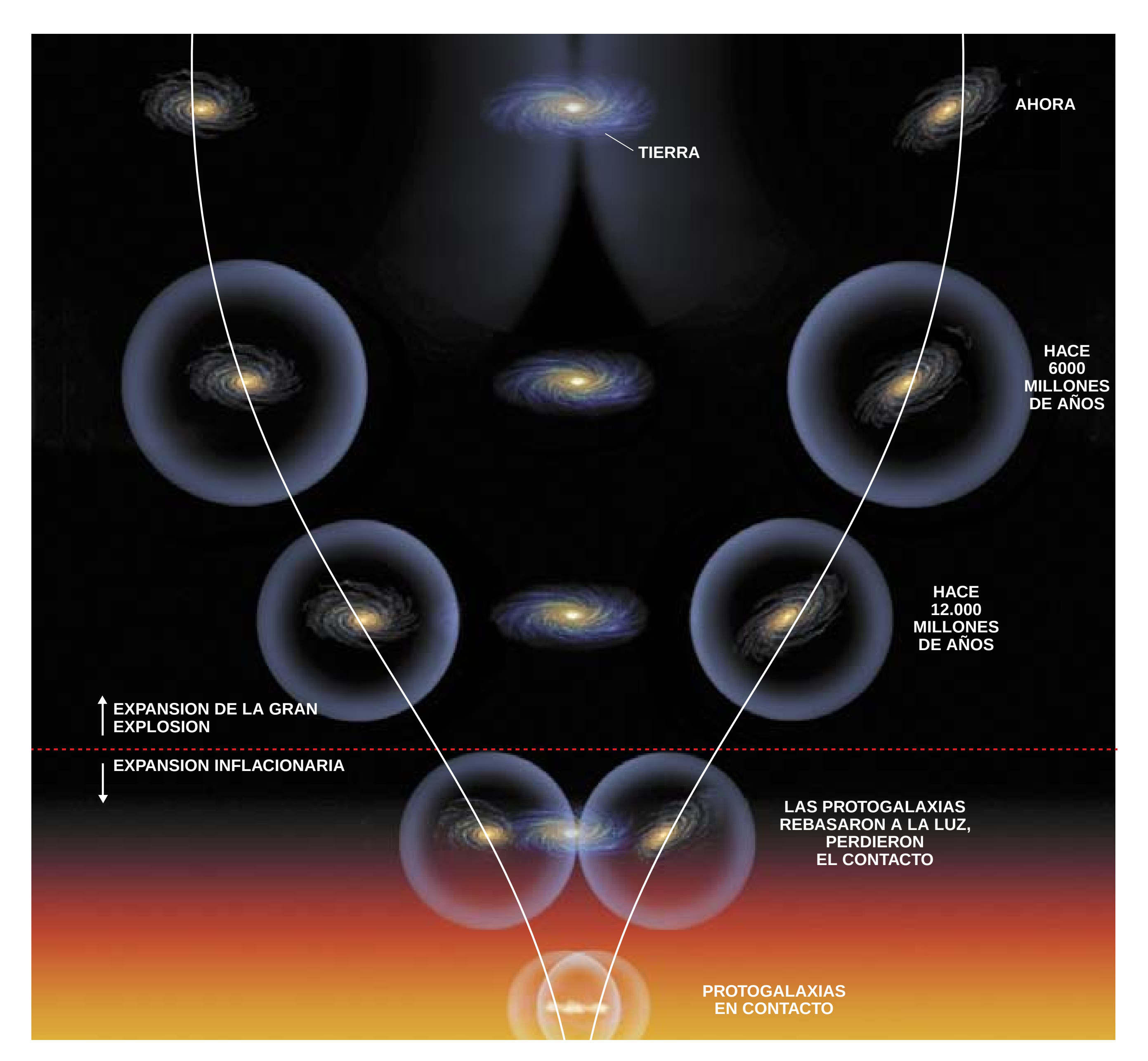
En un universo de burbuja, la vida presenta posibilidades interesantes. Por ejemplo, un observador externo podría pasar sin peligro del exterior hacia el interior de la burbuja. Pero una vez dentro, el observador (como nosotros) no podría salir nunca, pues para ello necesitaría viajar a velocidad mayor que la de la luz. Por otro lado, nuestro universo es uno más entre una infinidad de burbujas inmersas en un vasto y espumoso mar de vacío falso eternamente en expansión. ¿Qué pasaría si dos burbujas chocaran? Su encuentro desataría una explosión de proporciones cósmicas, destruyendo cuanto hubiere en el interior de las burbujas cerca del punto de impacto. Mas, por constituir la nucleación de burbujas un proceso rarísimo, esos cataclismos resultan improbables. Y aun cuando se produjera un episodio, quedaría exenta una parte sustancial de las burbujas. Para los observadores encerrados en las burbujas, aunque situados a una distancia prudencial, se trataría de un recalentamiento abrasador de una región del firmamento.

¿Podemos someter a prueba esta teoría? Explicar la razón de la homogeneidad del universo es buena cosa. Ahora bien, para dotar de validez a una teoría se requiere contrastar las predicciones cuantitativas de la misma con las observaciones reales. Los efectos específicos de inflación abierta fueron calculados en 1994 con aportaciones de los dos grupos que refinaron la teoría, así como de Bharat V. Ratra y P. James E. Peebles.

Los modelos inflacionarios antiguos y nuevos conceptos esbozan predicciones basadas en efectos cuánticos, que provocaron las diferencias sutiles en cuantía de inflación experimentada por distintos puntos del espacio. Cuando terminó el proceso inflacionario, quedó algo de energía en el campo inflatón; se convirtió en la radiación ordinaria, combustible de la expansión de la gran explosión subsiguiente. Puesto que la duración de la inflación divergía de lugar en lugar, también lo hacía la cantidad residual de energía y, por tanto, la densidad de radiación.

El fondo de radiación cósmica nos ofrece una instantánea de esas ondulaciones. En inflación abierta, el fondo recibe el influjo de las fluctuaciones que se desarrollan en el interior del universo, así como de las que se generaron fuera de la burbuja y se propagaron hacia dentro. Otras ondulaciones se ponen en movimiento provocadas por los desajustes en la nucleación de la burbuja. Esos efectos deberían mostrarse más patentes a escala cósmica, pues nos permiten mirar fuera de nuestro universo de burbuja. Uno de nosotros (Spergel), en colaboración con Marc Kamionkowski y Naoshi Suguiyama, fue más allá y advirtió que la inflación abierta debería tener otros efectos puramente geométricos.

Considerando el nivel de precisión de nuestro instrumental, las observaciones no pueden decantarse por una u otra de las dos teorías inflacionarias. El momento de la verdad vendrá con el lanzamiento previsto para finales del año próximo de la *Sonda de Anisotropías de Microondas* por la NASA. Una homóloga europea más avanzada, *Planck*, saldrá al espacio en el 2007. Efectuarán observaciones similares a las que realizó el satélite COBE, pero con una resolución mucho mayor. Deberían estar capacitadas para decidir qué teoría —si la constante cosmológica o la inflación abierta— es la correcta. Pudiera darse que ninguna de ellas satisficiese, en cuyo caso los investigadores tendrían que empezar de nuevo y buscar ideas nuevas sobre lo que aconteció en el universo primitivo.




**4. UNA GRAN PARADOJA COSMOLOGICA es la de la cuasiuniformidad del universo. En la expansión normal de la gran explosión, semejante regularidad resulta imposible (*parte superior del diagrama*). Hace miles de millones de años, dos galaxias instaladas en polos opuestos del firmamento empezaron a brillar. Aunque el universo se expandía, la luz pudo alcanzar otras galaxias y terminó por arribar a nuestra Vía Láctea. Desde sus telescopios, los humanos se percataron de que eran galaxias muy parecidas. Sin embargo, la luz de una galaxia no había llegado todavía a la otra, ni a la inversa. ¿Cómo, sin verse mutuamente, armonizaron su apariencia? La inflación (*parte inferior*) nos ayuda a responder. En las primeras fracciones de segundo de historia cósmica, los predecesores de las galaxias se tocaban. Entonces el universo se expandió a un ritmo acelerado, tirando de ellas y apartándolas más rápido que la luz. Desde entonces, las galaxias no se han vuelto a ver. Cuando la inflación terminó, la luz pudo alcanzarlas de nuevo. Tras miles de millones de años, las galaxias tornarán a establecer contacto.**

BIBLIOGRAFIA COMPLEMENTARIA

OPEN UNIVERSE FROM INFLATION. Martin Bucher, Alfred S. Goldhaber y Neil Turok en *Physical Review D*, vol. 52, n.º 6, págs. 3314-3337; 15 de septiembre, 1995. Reprint en xxx.lanñ.gov/abs/hep-ph/9411206 en la World Wide Web.

THE VERY FIRST LIGHT: THE TRUE INSIDE STORY OF THE SCIENTIFIC JOURNEY BACK TO THE DAWN OF THE UNIVERSE. J. Mather, con J. Boslough. BasicBooks, 1996.

THE INFLATIONARY UNIVERSE: THE QUEST FOR A NEW THEORY OF COSMIC ORIGINS. Alan H. Guth. Perseus Press, 1997.

Para más información sobre la Sonda de Anisotropías de Microondas y el satélite Planck consúltese la World Wide Web en map.gsfc.nasa.gov/ y astro.estec.esa.nl/SA-general/Projects/Planck/.

# PERFILES

Marguerite Holloway

## JAMES R. FLYNN: ¿Somos ahora más listos?

Los coeficientes intelectuales van subiendo, descubrió James R. Flynn. A él no parece importarle, seguro de que no somos más listos.

Ha terminado la clase, y se dirige como una flecha a su despacho para poner por escrito una ocurrencia feliz sobre Marx, el libre albedrío, el catolicismo y el invento de la máquina de vapor. Se le ocurrió al hilo de una pregunta en clase. En su clase de políticas, en la Universidad de Otago en Dunedin (Nueva Zelanda). Según me confiesa, las conversaciones improvisadas llevan a pensar de manera creativa y a concebir ideas nuevas.

Se muestra dispuesto a platicar sobre la tesis que le hizo famoso: la observación de que los coeficientes de inteligencia, tal y como se miden con algunas pruebas, han aumentado sin cesar desde comienzos de siglo.

Los hallazgos de Flynn, cuidadosamente documentados, han provocado una suerte de examen de conciencia entre psicólogos y sociólogos. Antes de que publicara sus investigaciones, allá por los años ochenta, las pruebas de inteligencia pasaban por mal momento. No faltaban voces críticas. Si bien, se aceptaban como un indicador, imperfecto ciertamente, de la agudeza y de diversas capacidades mentales o de su menoscabo. Pero tras el debate sobre el efecto Flynn —así se le llama— nada quedó igual que antes. Se polemiza sobre lo que realmente miden las pruebas, cuántas clases de inteligencia existen, si persisten las diferencias raciales y —caso de que el coeficiente intelectual esté realmente subiendo— por qué ocurre eso y cuáles son las consecuencias políticas y sociales.




*El activismo en pro de los derechos civiles le llevó a James R. Flynn al descubrimiento de que las puntuaciones de las pruebas de inteligencia aumentan con cada generación, un argumento de peso en favor de los factores ambientales, más que de los genéticos*

"Es un trabajo transformador", comenta Ulric Neisser, de la Universidad Cornell, editor de *The Rising Curve*. Este libro recoge las tesis de un simposio de 1996 de la Asociación Americana de Psicología, repasa el efecto Flynn y detalla las diversas explicaciones que se dan del mismo: mejor nutrición, atención familiar, extensión de la escolarización, mayor capacidad para resolver dificultades e impacto de los estímulos visuales y espaciales de un mundo lleno de televisores y videojuegos.

Flynn no privilegia ninguna de estas explicaciones en particular. Sentado en su despacho entre pilas de libros y artículos, tiene todo el aspecto de un Poseidón enjuto e irreverente: rizos grises, barba blanca, ojos azul claro y porte contestatario. Si el progreso de la inteligencia es real, "¿por qué no estamos experimentando un renacimiento sin precedentes en la historia humana?", cuestiona, casi con irritación. "¿Por qué no se repiten los días dorados de Atenas o del Renacimiento italiano?"

Sus propias convicciones humanísticas le indujeron a estudiar el coeficiente intelectual. Durante los años cincuenta fue militante activista de los derechos civiles en Chicago. Tras doctorarse, enseñó en la Universidad Oriental de Kentucky y presidió el Congreso de la Igualdad Racial en Richmond. "Como filósofo moral y político, mi interés principal radica en el uso de la razón y los datos que tenemos contra las ideologías totalitarias", explica. "Destacan entre éstas las ideologías racistas, pues el racismo ha sido uno de los principales adversarios contra los ideales igualitarios".

Su compromiso con los derechos civiles no le sirvió para encontrar trabajo en la universidad cuando era

joven. El y su mujer, Emily —cuya familia militaba en el Partido Comunista y, según Flynn, no fue ajena a la persecución— decidieron buscar una nación donde se encontraran a gusto. Se decidieron por Nueva Zelanda.

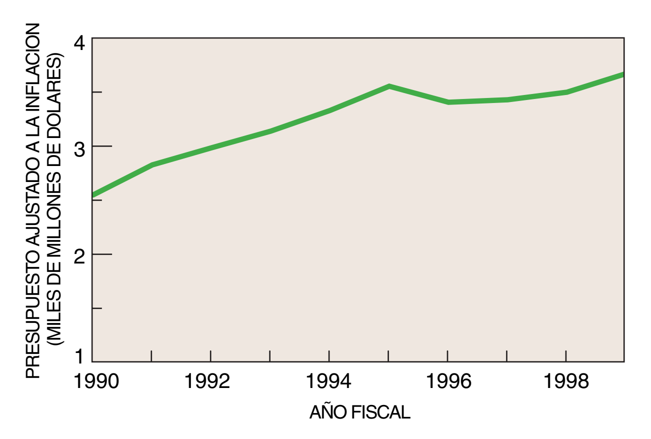
Una vez aposentados en su nuevo país y habiendo comenzado ya a educar a sus dos hijos, Flynn siguió luchando desde la distancia contra el racismo norteamericano. "Pensé que para impugnar eficazmente las ideas racistas tenía que examinar el debate sobre el coeficiente intelectual de las razas, la afirmación de que los negros, en promedio, son genéticamente inferiores." Se propuso refutar a Arthur R. Jensen, de la Universidad de California en Berkeley, uno de los principales adalides de ese punto de vista. En 1980 Flynn publicó *Race, IQ and Jensen*, y se desató el duelo. Decidió continuar con una breve monografía sobre las pruebas de inteligencia del ejército, porque tenía la corazonada de que los datos habían sido mal interpretados y que, de hecho, los reclutas negros progresaban más en cuanto al coeficiente intelectual que los blancos, una tendencia que apoyaría la convicción de Flynn de que el coeficiente intelectual estaba ligado más a los factores del entorno que a los genéticos.

Descubrió un error en el método de análisis de las encuestas militares. Conforme progresaba en su investigación, se dio cuenta de que Jensen y otros no aceptarían sus hallazgos so pretexto de que las pruebas del ejército estaban seriamente sesgadas por la educación, al revés que otras. Dicho de otra manera, la educación era un factor destacado en las puntuaciones. Como los reclutas negros tenían una mejor educación en los años 50 que en los años 20, cualquier ascenso en sus puntuaciones podría ser atribuido a la educación, y no a un incremento "real" del coeficiente de inteligencia.

No se arredró. Pensó que sería coser y cantar encontrar una prueba que midiera la inteligencia "genuina" y que se correlacionara con las pruebas del ejército, permitiéndole así usar los datos de estas últimas. No se encontró tal correlación, pero por el camino Flynn halló una mina de oro. Descubrió que algunas pruebas de inteligencia (en particular, las series Stanford-Binet y Wechsler) tenían versiones nuevas y antiguas y ambas se habían aplicado al mismo grupo de gente. En el caso de una de las pruebas Wechsler, por ejemplo, un mismo grupo de niños había sido examinado con las dos versiones. Y les salió mucho mejor la prueba de 1949 que la de 1974. Allá donde mirara, se daba cuenta de que los grupos salían mucho mejor parados con las pruebas antiguas. En 1984 Flynn dio a conocer que los americanos habían mejorado su coeficiente intelectual en 13,6 puntos en 46 años.

Aunque otros investigadores habían notado diferentes aspectos del fenómeno, siempre le habían quitado importancia. Pero Flynn no. "Creo que la razón principal es que, como yo no era psicólogo, no sabía lo que tenía que salir."

Los críticos, Jensen entre ellos, replicaron que la educación debía influir en las pruebas más de lo que se sospechaba. De modo que Flynn estudió las variaciones en las puntuaciones de una prueba llamada Matrices Progresivas de Raven, que mide lo que se llama *g* fluida: la resolución de problemas a bote pronto, que no está adulterada con influencias educacionales o culturales. En estas pruebas se emplean patrones en vez de elementos matemáticos o de palabras, por ejemplo. "Pueden realizarlo tanto los esquimales del Polo como los bosquímanos de Kalahari", señala Flynn. Sorprendentemente, resultó que los mayores aumentos se observaban en el Raven. Flynn observó que en 14 países (hoy ya tiene datos de al menos 20) el coeficiente intelectual estaba aumentando cada generación entre 5 y 25 puntos. "La hipótesis que mejor se ajusta a los resultados es que las pruebas no miden la inteligencia, sino que guardan correlación a través de un enlace causal débil con la inteligencia", escribió al publicar los datos. "Eso es lo que salió en el artículo de 1987 —dice— y claro, fue como soltar al zorro en el gallinero".




*Las puntuaciones de las pruebas de inteligencia han subido sin cesar durante más de 50 años en todo el mundo*

Flynn acaba de descubrir otro aumento raudo y desconcertante en las puntuaciones de uno de los ejercicios del Wechsler, uno que sirve para medir la capacidad verbal. Antes de su nuevo hallazgo, parecía verosímil que las puntuaciones del Raven aumentaran debido a la difusión de los videojuegos y la universalización del uso del ordenador. Pero ahora el misterio se ha vuelto impenetrable.

A pesar de dos décadas de debates con Jensen, manifiesta sentir la consideración más profunda hacia este académico y su escuela. "En la izquierda existe la tentación de no querer mirar los datos", subraya. "Si a Arthur Jensen le asiste la razón, aquí tenemos una verdad significativa sobre el mundo real a la cual todos debemos adaptarnos." Flynn dice que quiere que reinen los principios del igualitarismo humanitario "allí donde tenga el valor de enfrentarme con todos los aspectos del mundo real. Y si uno de esos aspectos es que los negros —en promedio, no individualmente— son genéticamente inferiores en alguna clase de inteligencia que pasa factura en la era de la informática, haríamos bien en saberlo".

La siguiente pregunta, claro está, es si cree que existan tales diferencias. "!En absoluto!"

Además de su trabajo sobre el coeficiente intelectual, batalla en otros frentes. Decepcionado por la derrotista marcha de Nueva Zelanda hacia el capitalismo puro, ha procurado frenar la ola presentándose en las elecciones parlamentarias. Tres veces ha sido candidato y tres veces lo han vapuleado. La última, en 1996. Se presentó por el Partido de la Alianza: "El único partido de Nueva Zelanda que todavía cree en los impuestos como medio de redistribuir la riqueza y defiende una sanidad y una educación para todos".

Acaba de terminar un quinto libro, *How to Defend Human Ideals*, al que ha dedicado bastantes años. "Probablemente a nadie le importe; la gente se interesa mucho menos en las contribuciones fundamentales que en las espectaculares", se lamenta.

# Ciencia y sociedad

## Transmisión sináptica

### *El receptor de cainato*

Cuando la terminación de un axón de una célula nerviosa hace contacto con el cuerpo o dendrita de otra neurona se establece una sinapsis. Esta puede ser excitadora o inhibidora. La mayoría de las sinapsis excitadoras del sistema nervioso central (SNC) utilizan al aminoácido glutamato como neurotransmisor.

Liberado en la hendidura sináptica, el glutamato interactúa con diversos receptores específicos y da lugar a un variado repertorio de respuestas excitadoras en las neuronas. Se distinguen dos grandes familias de receptores específicos de glutamato, o glutamatérgicos: los metabotrópicos (receptores acoplados a sistemas intermediarios, de segundos mensajeros) y los ionotrópicos (canales iónicos activados por glutamato). Estos últimos a su vez se subdividen en los tipo NMDA, AMPA y cainato, que toman el nombre de la sustancia agonista preferencial que los activa. (Llámanse agonistas las moléculas que al unirse a un receptor incrementan la actividad del mismo.)

Se conocen los pormenores de la participación de los receptores tipo AMPA y NMDA en la transmisión sináptica excitadora. Por el contrario, la función de los receptores de cainato en la sinapsis ha permanecido como una incógnita. Por un lado, no es fácil distinguir sus efectos de los inducidos por los receptores AMPA. Se sabe, por otro lado, que el cainato es un poderoso agente neuroexcitante, capaz de inducir epilepsia y degeneración neuronal selectiva. Estos hechos han estimulado durante años la búsqueda de una función de los receptores de cainato.

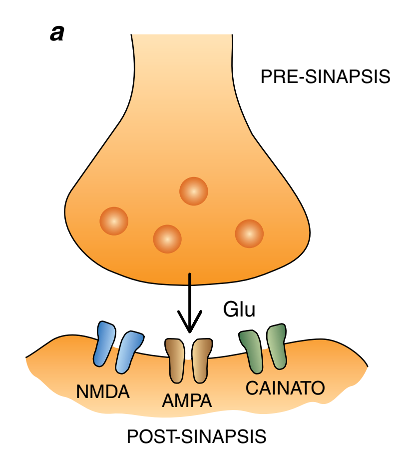
La aparición reciente de antagonistas específicos que permiten distinguir entre los receptores AMPA y cainato, así como la posibilidad de alterar selectivamente la expresión de receptores por manipulación genética en ratones, han comenzado a dar frutos que contribuyen a comprender la función de los receptores de cainato en la transmisión sináptica.

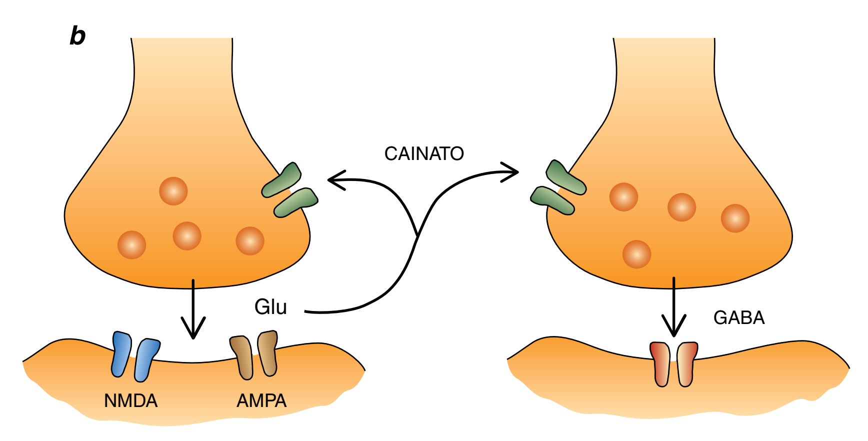
Al igual que los otros receptores glutamatérgicos, los receptores de cainato estarían formados por un conjunto de subunidades distintas que configuran un canal catiónico. Se han clonado cinco subunidades del heterómero que poseen una elevada afinidad por el cainato: GluR5-6-7, KA-1 y KA-2. Pero seguimos sin conocer la composición exacta de los receptores en las neuronas.

Una primera sospecha sobre la participación de los receptores de cainato en la transmisión sináptica proviene de una doble observación, a saber, su particular localización en zonas blanco de vías excitadoras glutamatérgicas y su ubicación presináptica y postsináptica a tenor de las sinapsis consideradas. Concordantes con estos datos anatómicos, se han aportado pruebas experimentales que apoyan dos modelos, no excluyentes, sobre la función sináptica de los receptores de cainato. De acuerdo con el primer modelo, los receptores de cainato pueden ejercer un efecto postsináptico directo y mediar en las corrientes excitadoras; en el marco del segundo, los receptores pueden modular la liberación de neurotransmisores mediante un efecto presináptico.

Gracias al bloqueo selectivo de los receptores tipo AMPA y NMDA, hemos podido poner de manifiesto que el glutamato activa receptores postsinápticos de cainato. Es decir, este neurotransmisor, además de inducir respuestas postsinápticas excitadoras por activación de los receptores ionotrópicos tipo AMPA y NMDA, genera en la misma sinapsis corrientes excitadoras mediadas por receptores de cainato. Se trata de un fenómeno que se observa en las neuronas piramidales de la región CA3 del hipocampo, uno de los primeros sitios reconocidos en el sistema nervioso central con una unión de alta afinidad para el cainato y con una singular vulnerabilidad ante su acción neurotóxica.

Las corrientes sinápticas excitadoras mediadas por receptores de cainato están ausentes en ratones mutantes a los que se les alteró la subunidad GluR6. Se confirmaba así la participación de estos receptores en la transmisión sináptica. Del ensayo se desprendía, además, que la subunidad formaba parte del receptor. Lo que a su vez explicaba la mayor resistencia que los ratones mutantes oponían a




*Dos modos de acción de los receptores de cainato en la sinapsis de acuerdo con su localización postsináptica* (a) *o presináptica* (b)

la inducción de epilepsia y efectos neurotóxicos por cainato.

Entre los efectos presinápticos de los receptores de cainato se ha propuesto que la activación de éste inhibiría la liberación de glutamato en el hipocampo. Además de esta acción sobre la sinapsis excitadora, el glutamato liberado podría difundir hacia otras sinapsis locales y ejercer efectos distintos. Así, se ha postulado que la activación de receptores de cainato localizados en terminales sinápticas inhibidoras del hipocampo de roedores disminuiría la liberación del neurotransmisor GABA, que es un inhibidor de la sinapsis en vertebrados superiores. De tal modo, los receptores de cainato también influirían sobre la excitabilidad neuronal indirectamente al reducir la actividad inhibidora de naturaleza GABA y podrían explicar, en parte, la acción neuroexcitante del cainato.

Los estudios actuales permiten establecer una participación de los receptores de cainato en la transmisión sináptica, ya sea directamente, mediando una actividad excitadora en la neurona postsináptica, o bien indirectamente, modulando la liberación de los neurotransmisores GABA y glutamato. Es conocido que la actividad excitadora glutamatérgica es clave en fenómenos tales como la plasticidad sináptica, la sinaptogénesis y la conformación de circuitos neurales durante el desarrollo. Por otra parte, una excitación excesiva por glutamato puede inducir actividad epiléptica y muerte neuronal. Las investigaciones futuras profundizarán seguramente sobre la función de los receptores de cainato en distintas sinapsis del SNC y sobre su contribución en los mecanismos de la epilepsia.

PABLO E. CASTILLO
Depto. de Fisiología
Universidad de la República,
Montevideo (Uruguay)

## Microelectrónica

### *Implantación iónica*

Para dotar a una pieza de mayor dureza o mejor resistencia contra la corrosión, los metalúrgicos acuden a veces a la implantación iónica. Pero no se restringe ahí su ámbito de aplicación. Por su versatilidad a la hora de controlar las propiedades de un semiconductor, la implantación iónica es una técnica que experimenta una fuerte expansión en microelectrónica. Consiste en la introducción de átomos de un cierto elemento dentro de un sustrato sólido; para lograrlo, se bombardea el sustrato con un haz de iones acelerados del elemento a implantar, en condiciones de alto vacío. Esta técnica surgió de los avances experimentados en los campos de la ingeniería de aceleradores de partículas y separadores de isótopos.



*1. Imagen de un implantador iónico de los utilizados habitualmente en microelectrónica*

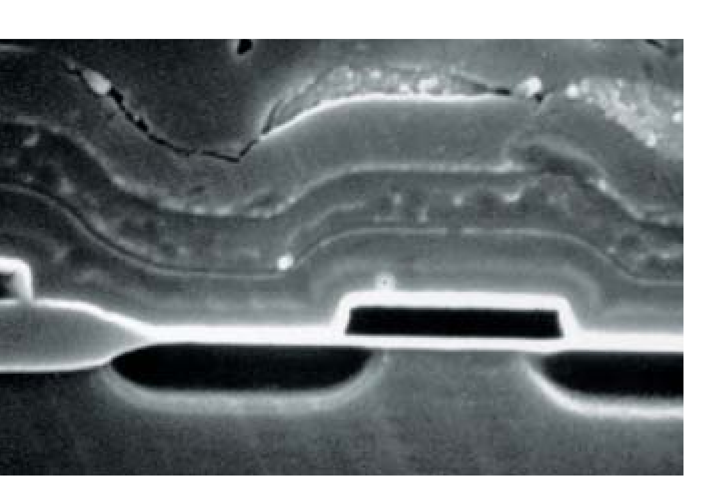
Se requieren de 15 a 20 etapas de implantación iónica para fabricar un microprocesador de última generación o una memoria RAM de 4 megabits, componentes fundamentales de los ordenadores personales. Mediante la introducción de una cantidad adecuada de impurezas dopantes podemos cambiar la conductividad eléctrica de los semiconductores en varios ordenes de magnitud.

Un implantador iónico consta de una fuente de iones, un electroimán de selección del ion a implantar, una columna aceleradora y sistemas de focalización, deflexión y barrido del haz iónico. Al penetrar en el sólido, los iones acelerados van cediendo energía en sus interacciones con núcleos de los átomos del blanco y en las interacciones con electrones libres. Se da, pues, un doble frenado: el nuclear y el electrónico. El frenado nuclear fue cuantificado por J. Lindhard, M. Scharff y H. E. Schiott (teoría LSS) usando la mecánica clásica de colisiones binarias y el potencial interatómico apantallado de Thomas-Fermi. En un modelo de gas de electrones, el frenado electrónico es proporcional a la velocidad del ion.

Considerando ambos tipos de frenado, se puede determinar la trayectoria (R) del ion dentro del sólido. La magnitud más interesante es, sin embargo, la proyección de R sobre la normal a la superficie o rango proyectado ($R_p$). Conviene destacar que el frenado es un fenómeno aleatorio y su cálculo, probabilístico.

En materiales amorfos, la distribución de iones implantados se aproxima a una distribución gaussiana. En materiales cristalinos los iones energéticos penetran más profundamente en los sustratos si se implantan en direcciones cristalográficas de bajo índice (fenómeno de acanalamiento).

El proceso de implantación, cualquiera que sea, daña siempre el material bombardeado. La razón estriba en que, si se transfiere energía suficiente a los átomos del sustrato por medio de choques nucleares se producen desplazamientos de sus posiciones en la red. Después de un proceso de implantación iónica se requiere un recocido a alta temperatura, entre 900 y 1100 ºC para el silicio, para reparar lo estropeado, favorecer que los iones implantados se instalen en el entramado reticular y provocar su activación eléctrica.



*2. Fotografía obtenida por microscopía electrónica, donde se pueden apreciar algunas de las regiones implantadas en un transistor MOS; este dispositivo corresponde a un circuito integrado fabricado en el CNM*

Para especificar la implantación iónica se emplean dos parámetros: la energía y la dosis (átomos implantados por unidad de superficie). La profundidad de penetración, proporcional a la energía, es mayor en los iones ligeros. Los valores de energía habituales varían entre 20 y 200 kiloelectronvolt. Para estas energías, y considerando las impurezas comunes en el caso del silicio (boro, fósforo, arsénico), las distribuciones de impurezas no alcanzan una profundidad superior a 1 micrometro. Por lo que respecta al segundo parámetro, el de las dosis, los valores típicos están entre $10^{12}$ y $10^{16}$ átomos por centímetro cuadrado.

Entre las ventajas de la implantación iónica merece resaltarse el control del perfil de concentración de las impurezas introducidas. Además, pueden emplearse materiales muy dispares (por ejemplo, resina fotosensible) como máscaras para implantar selectivamente el dopante. Por fin, se alcanza una buena homogeneidad y repetitividad. Entre las desventajas de la implantación iónica no podemos silenciar la necesidad de un tratamiento térmico subsiguiente a la implantación propiamente dicha y que es un proceso muy caro.

En el ámbito de la técnica de semiconductores, y en particular CMOS-LSI (C*omplementary* M*etal* O*xide* S*emiconductor*-L*arge* S*cale* I*ntegration*), se han encontrado numerosas aplicaciones a la implantación iónica: 1) implantación de campo, que evita la aparición de transistores MOS parásitos de campo; 2) implantación del pozo, donde se realiza el transistor complementario; 3) implantación de ajuste de la tensión umbral de los transistores MOS de canal P y de canal N; 4) implantación para obtener las regiones de fuente y drenador de los transistores MOS de canal P y de canal N; 5) implantación del silicio policristalino utilizado como puerta; 6) implantación de drenador ligeramente dopado, o LDD (L*ightly* D*oped* D*rain*) para reducir el pico de campo eléctrico en el extremo del canal y, de esta manera, disminuir la generación de portadores calientes, y 7) implantación para aumentar la tensión de perforación.

Con la llegada de la técnica VLSI (V*ery* L*arge* S*cale* I*ntegration*) se plantean nuevas necesidades y problemas en el campo de la implantación iónica. Su solución pasa por la innovación del proceso y de los equipos.

Entre las nuevas áreas de interés destacan las implantaciones a muy baja energía (1 kiloelectronvolt) y alta dosis para la obtención de uniones poco profundas, consecuencia de la reducción de las dimensiones superficiales de los transistores MOS con el fin de conseguir mayores densidades de integración. Para la obtención de pozos con un perfil retrógrado y de capas enterradas se recurre a la implantación iónica de alta energía (por encima del megaelectronvolt). La ventaja principal de esta implantación iónica a altas energías reside en la reducción del número de etapas de fabricación, con la consiguiente reducción del coste.

Se muestra prometedora también la implantación iónica con $O_2$ a dosis altas para crear una capa enterrada aislante de $SiO_2$. Este proceso de separación mejora la velocidad de conmutación de los dispositivos. Cabe resaltar, asimismo, la implantación iónica con haz paralelo y con grandes ángulos de incidencia, cuya importancia crece al aumentar el diámetro de las obleas de silicio. Por último, la implantación iónica con microhaz focalizado permite la implantación directa sin necesidad de enmascaramiento.

JOSEP MONTSERRAT
Centro Nacional
de Microelectrónica (CNM),
CSIC, Bellaterra, Barcelona

# Muestreo

## *Grupos disjuntos y no disjuntos*

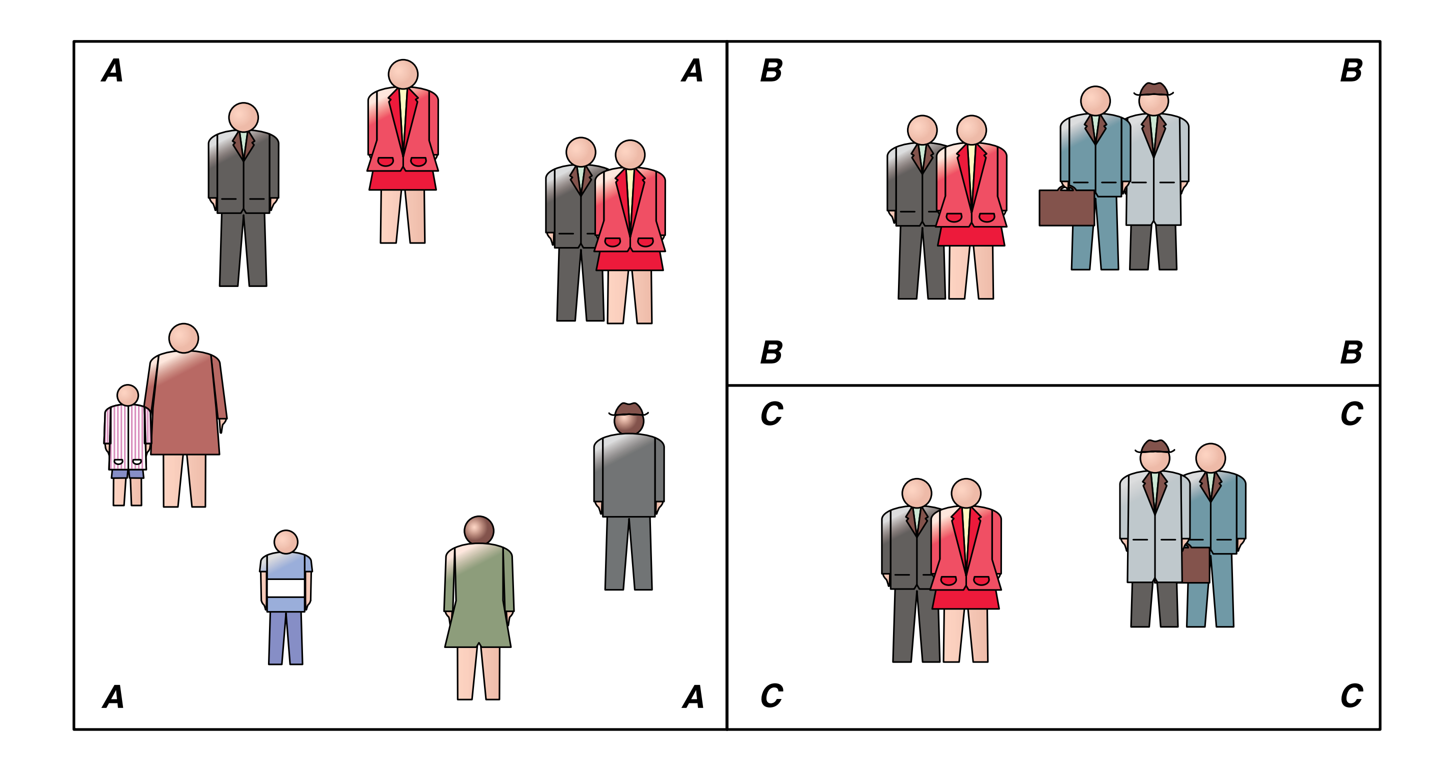
El estudio de características relativas a una población, entendiendo por ésta a un conjunto de unidades en las que cada una contribuye con su aportación individual al global que se desea conocer o estimar, ha sido objeto de interés permanente. Pero ha sido a lo largo de nuestro siglo cuando se han desarrollado las mejores teorías estadísticas que permiten descubrir tales características globales, con sólo observar una parte o muestra (seleccionada aleatoriamente y de modo controlado) de la población.

Sea, por ejemplo, la intención de voto por un partido político en campaña electoral. Aunque podría considerarse definido en el momento en el que se le pregunta y responde el encuestado, no coincide necesariamente con la decisión que tomará éste el día de la votación. Otro ejemplo, la declaración fiscal de la renta personal y del patrimonio. Los datos del declarante se recogen en un impreso normalizado; cualquier dato proporcionado está definido y es concreto.

En el primer ejemplo, el dato de intención de voto es una respuesta libre; no está, pues, determinado de antemano ni es fijo para cada elector. En el segundo ejemplo el dato consignado y firmado por el declarante tiene un rango legal que exige su veracidad frente a posibles fraudes; es decir, en este caso son fijas y definidas todas las cantidades consignadas.

Tales hipótesis de trabajo admiten una formalización basada en poblaciones finitas. Además, de cada unidad se puede observar su dato fijo para la característica que se desea conocer del conjunto global.

Entre los métodos de muestreo los de agrupación disjunta tienen una historia más dilatada. Tschuprow y Tischendorff en los años veinte, Neyman en los treinta y Glasser y otros expertos, más recientemente, impulsaron el progreso de la asignación muestral eficiente y optimización de la estratificación. Llámase estratificación a la clasificación disjunta de las unidades de la población en estratos, entendiendo por éstos a grupos de unidades homogéneas o similares según la característica en estudio, pero heterogéneos o dispares dos a dos.




*En el gráfico superior la población de interés está formada por personas que individualmente están incluidas en un solo grupo o rectángulo cuyos "vértices" se identifican por la letra correspondiente en el ángulo interior de ellos. Son grupos disjuntos. El gráfico inferior corresponde a grupos no disjuntos. Cada rectángulo A, B y C se ha delimitado por su letra en el ángulo interior de sus vértices. Así el rectángulo A incluye un grupo de 12 personas, el B, 8 personas, y las personas que están en A y B simultáneamente son 3*

Otro método de clasificación disjunta de la población es la conglomeración, así se denomina la clasificación de las unidades en conglomerados. Menos estudiada desde el punto de vista de la precisión de que es capaz, se la ha venido confinando al campo administrativo donde interesa aprovechar recursos disponibles como los censos de población que clasifican a los habitantes en secciones censales. En las técnicas muestrales aplicadas a estudios socioeconómicos, cada sección censal suele tratarse como un conglomerado o clase disjunta de personas con características heterogéneas en su clase; las características son homogéneas entre sí, para que las estimaciones basadas en el muestreo por conglomerados básico resulten eficientes.

Un grupo-estrato difiere de un grupo-conglomerado en que el primero se muestrea (parcial o totalmente) con seguridad, mientras que el segundo sólo se muestrea con cierta probabilidad habitualmente no nula.

La teoría de muestras considera unidades abstractas en su razonamiento matemático; en la práctica pueden tomar cuerpo en personas, países, especies, objetos, cultivos agrarios y un sinfín de posibles variables o características de interés que justifiquen su aplicación. Aspectos sociales, producciones agrarias e industriales, económicos y demográficos son algunos de sus campos de interés clásicos. Ahora se aplican también a la biodiversidad y la ecología.

Las técnicas de muestreo estratificado y por conglomerados han sido generalizadas en los últimos diez años por S. Singh y por el autor de esta nota en colaboración con el profesor L.N. Sahoo. En estos trabajos de investigación, los conglomerados o estratos pueden superponerse, es decir, puede haber alguna o algunas unidades de la población que pertenezcan a dos o más grupos.

Ejemplo: en la población de propietarios de inmuebles en países miembros de la Comunidad Europea, es posible que algún propietario pueda poseer inmuebles en varios países comunitarios. En este caso, los grupos de propietarios europeos por países no son todos disjuntos dos a dos.

Hace pocos años el estudio de métodos muestrales con propiedades y cualidades estadísticas razonables que pudieran estimar características globales a partir de grupos superpuestos o solapados constituía un campo virgen. Hoy se ha avanzado bastante en la fundamentación de metodologías que resuelven estas cuestiones, contándose además con las posibilidades de almacenamiento masivo de datos, programación y cálculo que la informática actual ofrece.

Las teorías basadas en presupuestos proporcionales a los tamaños muestrales, aunque no han perdido vigencia, se han visto superados (especialmente en paneles y bases de datos informatizados, pues el coste muestral es muy inferior al de programación y diseño) cuando se trata de lograr una precisión suficiente para la toma de decisiones que reviertan en el bien común.

MARIANO RUIZ ESPEJO
Facultad de Ciencias Económicas, UNED, Madrid

## *Escherichia coli*

### *Genoma al desnudo*

*Escherichia coli* K-12 ocupa un lugar singular en biología. Se trata de la célula autónoma que mejor conocemos. Esta cepa bacteriana, aislada en 1922 de un paciente diftérico, se convirtió en modelo de estudio experimental a raíz del descubrimiento de su función sexual en 1946.

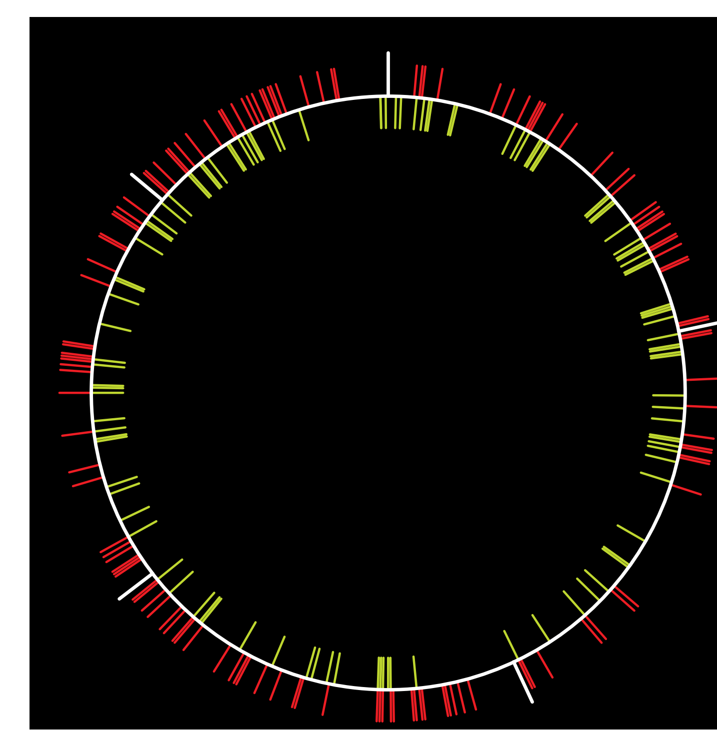
En ese organismo se incubó el nacimiento de la biología molecular. Y en él se operaron los descubrimientos fundamentales para entender la lógica molecular de los organismos. El 16 de febrero de 1997 se depositó en las bases de datos la secuencia completa de su cromosoma: 4.639.221 nucleótidos.

El trabajo de secuenciación desde su concepción llevó más de 5 años en la Universidad de Wisconsin en Madison, en el laboratorio de uno de los autores (Fred Blattner). Hubo que crear bibliotecas de fragmentos traslapados del genoma; primero, con secuencias del orden de 20.000 pares de bases clonadas en vectores del bacteriófago lambda; luego, con fragmentos diez veces mayores aislados directamente del genoma. Además de recurrir a otras técnicas propias de la ingeniería genética y de secuenciación automática.

Aunque en el transcurso de años anteriores se habían secuenciado ya cerca del 25 % de los genes, para evitar errores, nos empeñamos en secuenciar de nuevo el genoma entero. Se supone que un 90 % del genoma determina alguna proteína; el restante 10 % se encargaría de funciones reguladoras. No debe sorprendernos tamaña densidad de codificación, pues se trata de bacterias, y éstas, a diferencia de organismos superiores, poseen genomas compactos. La levadura, cuyo genoma cuadruplica el bacteriano, apenas si presenta un 50 % más en número de genes.

Del análisis de la secuencia se deriva el descubrimiento de la redundancia de ciertos genes (duplicados), el descubrimiento de capacidades metabólicas desconocidas en *E. coli* y la similaridad entre diversos genes de *E. coli* que dan lugar a familias de proteínas. La aplicación de métodos computacionales al genoma descifrado nos permite barruntar la red de interacciones reguladoras en que se halla involucrado.

Con los genomas bacterianos a nuestro alcance, la microbiología se dispone a dar el gran salto adelante para entender el origen de la vida,



*428 genes con mecanismos de regulación conocida en el genoma de* E. coli

los mecanismos de la evolución y las posibilidades de la ingeniería genética. De entrada se ha identificado la función bioquímica de cada gen. Hemos asimismo ubicado las señales más pequeñas, pero no por ello menos significativas, indicativas de las interacciones reguladoras que determinan la activación e inhibición de los genes ante distintas condiciones ambientales.

Queda todavía mucho por desentrañar. Las funciones de casi la mitad del genoma. En su desciframiento nos habrá de ayudar la fracción ya entendida. Si tuviéramos en un disco compacto los 4.639.221 símbolos ordenados del genoma de *E. coli*, comprobaríamos que no contamos con un decodificador que nos dé a la perfección el sentido de esta información, a diferencia de un disco compacto que porte el código de una sinfonía de Beethoven. La construcción de estos decodificadores de genomas es el reto para los próximas décadas.

Por otro lado, percibimos ya las limitaciones de los métodos computacionales y matemáticos para predecir el conocimiento biológico, donde abundan las excepciones y reina la complejidad. El hombre tendrá que esforzarse por idear métodos capaces de estudiar y formalizar sistemas complejos, heterogéneos, donde los elementos interaccionan entre sí. El modelo puede hallarlo en un genoma entero.

JULIO COLLADO-VIDES
Centro de Investigación sobre Fijación de Nitrógeno.
U.N.A.M. México
FRED BLATTNER
Laboratorio de Genética
Universidad de Wisconsin en Madison

## Domesticación del perro

### *Relaciones entre perro y lobo*

Suelen admitirse tres grandes explosiones demográficas en la historia de la humanidad. La primera se produjo con el descubrimiento del fuego, que le permitió extender su zona de habitación hacia latitudes más frías. La segunda se asocia con la domesticación de plantas y animales. La domesticación independiza de la aleatoriedad del entorno. Asegura el alimento a lo largo del año, sin verse obligado a migrar tras las presas. El nómada se tornó sedentario. Las ciudades crecieron, los modos de vida se diversificaron y se fueron colonizando nuevas tierras. Transcurridos varios miles de años, el advenimiento de la revolución industrial puso las bases de la tercera explosión demográfica, de la que somos protagonistas.

El perro parece ser la primera especie animal que el hombre domesticó. En yacimientos arqueológicos de Irán, Israel y Alemania se han encontrado los restos más antiguos, que datan de hace unos 14.000 años. A menudo aparecen enterrados juntos hombres y perros. Según el registro arqueológico, bastaron dos o tres mil años para la propagación del can por todo el mundo. Todas las demás especies domésticas aparecieron con posterioridad. Quizá la domesticación del perro aportó la habilidad técnica para la domesticación de los otros.

La domesticación del perro parece haber desempeñado una función muy importante en la evolución humana. Pero, ¿cuál es el origen del perro? ¿Cómo, cuándo y dónde se produjo esta domesticación? La variabilidad morfológica del perro es mucho mayor que la de cualquier otra especie doméstica. Las diferencias que observamos entre un galgo afgano, un bulldog, un sabueso o un podenco son mayores que las observadas entre las demás especies de cánidos. ¿Cuál es el origen de tal variabilidad? La comparación de las secuencias de ADN mitocondrial nos puede dar la respuesta.

Las mutaciones en las secuencias de ADN se acumulan con el tiempo. Dos especies que compartan un antepasado cercano guardarán, pues, mayor similaridad en sus secuencias de ADN que especies que han divergido mucho antes. Por tanto, si comparamos las secuencias de ADN de genes determinados, podremos reconstruir el proceso evolutivo que ha dado lugar a la aparición de una especie.

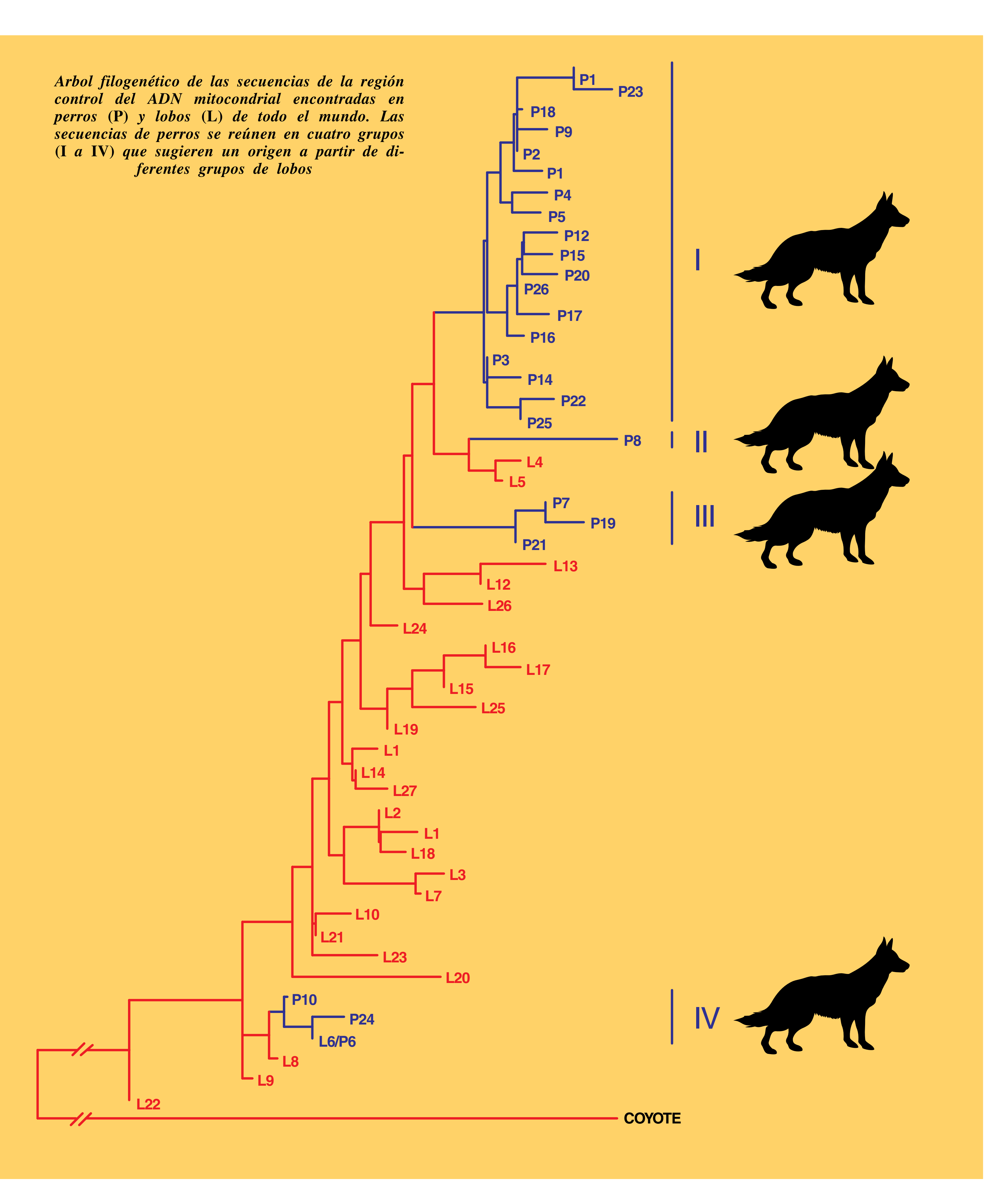
En colaboración con mis colegas de la Universidad de California en Los Angeles, del Instituto Real de Tecnología de Estocolmo, de la Universidad A&M de Texas y de la Universidad Brigham Young de Utah, hemos secuenciado un fragmento de la región control del ADN mitocondrial de coyotes, tres especies de chacales, perros pertenecientes a una gran variedad de razas y lobos de todo el mundo.

Las secuencias de ADN de los perros guardaban un parecido mucho más estrecho (cuando no eran incluso idénticas) con las de cualquier lobo que con las secuencias nucleotídicas de los otros cánidos. Ello significa que los perros proceden de lobos domesticados. Esta conclusión contradice las clásicas teorías de Darwin o Lorenz que sugerían que los antecesores del perro eran dos especies, el lobo y el chacal dorado.

Al reconstruir el árbol filogenético con las diferentes secuencias encontradas en perros y lobos se observa que los genotipos de los perros se reúnen en cuatro grupos. Dicho de otro modo, el ADN de los perros se originó a partir de cuatro grupos de lobos, por lo menos. ¿Cómo? Tal vez ocurrió en virtud de múltiples domesticaciones en diferentes lugares y tiempos o se produjo mediante una sola domesticación con la posterior introducción de lobas dentro de las poblaciones de cánidos domésticos.

Por otro lado, la variabilidad de uno de los grupos de perros (*I en la figura*) resulta sorprendentemente elevada si la domesticación aconteció hace sólo 14.000 años. Teniendo en cuenta la divergencia entre las secuencias de lobos y coyotes, especies que según el registro fósil están separadas desde hace más de un millón de años, podemos estimar la tasa de evolución del gen que estamos estudiando. Tras comprobar que en los perros no se ha dado ningún cambio significativo en la velocidad de evolución con respecto a los lobos, podemos estimar el tiempo necesario para generar la diversidad observada en el grupo I a partir de una única secuencia inicial: ¡135.000 años!

Aunque se trata de una estimación sujeta a amplios márgenes de confianza, la información genética es contundente. De la misma se desprende que la domesticación del perro es un fenómeno mucho más antiguo que los 14.000 años que se deducían de los registros arqueológicos. Quizás ambas fuentes de información —la genética y la arqueológica— discrepan tanto porque la morfología de los primeros perros se parecía a la de los lobos, por cuya razón sus restos fósiles resultan difíciles de identificar en yacimientos; o tal vez la discrepancia se deba a que la domesticación se produjo en zonas apenas exploradas por la arqueología.




*Arbol filogenético de las secuencias de la región control del ADN mitocondrial encontradas en perros (P) y lobos (L) de todo el mundo. Las secuencias de perros se reúnen en cuatro grupos (I a IV) que sugieren un origen a partir de diferentes grupos de lobos*

Si bien el perro deriva de una sola especie de cánido silvestre, su origen a partir de diferentes grupos de lobos y la antigüedad de la domesticación nos ayudan a desentrañar las causas de la variabilidad de perros existente en la actualidad.

No obstante, permanece envuelto en el misterio el modo en que se produjo la asociación entre perros y lobos, así como la permanente separación ulterior de sus poblaciones. El estudio del proceso de la domesticación del perro puede traer consigo nueva luz sobre la vida de las antiguas comunidades nómadas y la historia, paralela, del hombre y el perro.

CARLES VILÀ
Departamento de Biología Evolutiva
Universidad de Uppsala

## Quinasas de quinasas

### *Comunicación entre la membrana y el núcleo celular*

La adaptación de una célula a los cambios de las condiciones del medio que la rodea entraña la transmisión de instrucciones desde la membrana hasta el núcleo celular para poder generar las respuestas adecuadas.

La fosforilación de proteínas constituye un mecanismo clave en el control de la actividad de muchas enzimas. De dicha fosforilación se encargan unas enzimas llamadas proteína quinasas. Las proteína quinasas y las proteína fosfatasas (responsables de la defosforilación) intervienen en la regulación y coordinación de muchos aspectos esenciales en la vida celular, vale decir, en la expresión génica, el crecimiento y división celular o en diferentes aspectos del metabolismo. Ello significa que el control de la actividad de estas proteínas quinasas y fosfatasas resulta decisivo para ejecutar correctamente tales funciones.

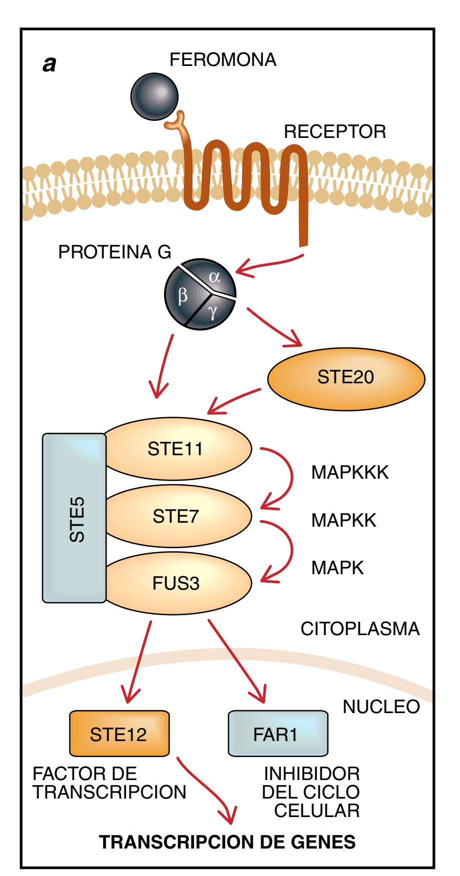
La actividad de las proteína quinasas puede regularse de muchas maneras. Las hay que quedan controladas mediante la unión a subunidades o dominios reguladores, así como a través de la interacción con inhibidores. Otro mecanismo común en la regulación de las quinasas es mediante su fosforilación.

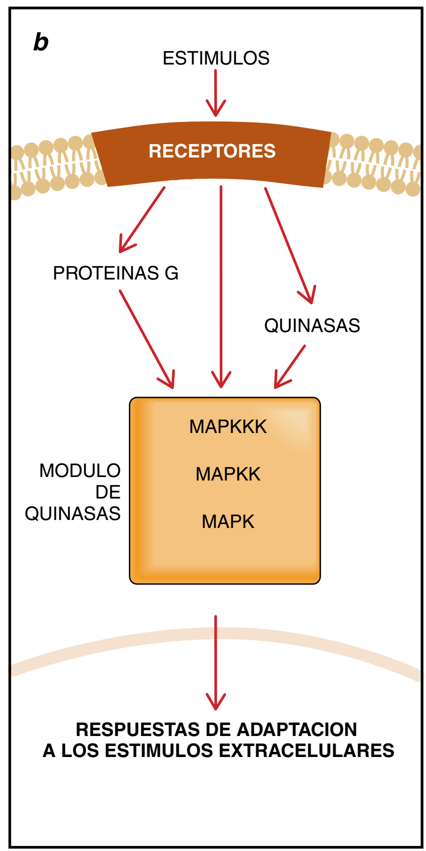
Con la cristalización y determinación de la estructura tridimensional de estas quinasas se ha descubierto un aspecto clave de la fosforilación. Sabemos ahora que determinadas quinasas pueden fosforilarse en unos residuos situados en el centro del dominio catalítico de la quinasa. El segmento activador, así se llama, actúa como si fuese un labio flexible y su fosforilación provoca un cambio de su conformación que libera el sitio activo de la quinasa, previamente bloqueado, permitiendo el acceso del substrato para su posterior catálisis.

Varias son las quinasas que pueden regularse por vía de su fosforilación; entre ellas, las MAP quinasas ("Mitogen Activated Protein"). Estas enzimas se fosforilan rápidamente en respuesta a estímulos extracelulares. Una vez las MAP quinasas han quedado fosforiladas en el segmento de activación (generalmente en treonina y tirosina), se tornan catalíticamente activas —capaces de fosforilar otras proteínas— y pueden trasladarse al núcleo de la célula.

En el núcleo celular las MAP quinasas fosforilan factores de transcripción específicos y modifican así su actividad, con lo que provocan un cambio en el patrón de expresión de un conjunto de genes que permitirán a la célula adaptarse a las nuevas condiciones del entorno. Las MAP quinasas, sin embargo, son sólo un eslabón más de la cascada de quinasas.

Por cascada de quinasas se entiende la intervención sucesiva de tres quinasas relacionadas. Una MAP quinasa (MAPK) es activada por una MAPK quinasa (MAPKK), activada a su vez y también mediante fosforilación por una MAPKK quinasa (MAPKKK). Estos grupos de quinasas se han encontrado en todas las células eucariotas y forman módulos




*Los módulos de quinasas intervienen en la comunicación entre la membrana celular y el núcleo. La levadura* S. cerevisiae *utiliza una vía de MAP quinasas compuesta por STE11, STE7 y FUS3 para responder a la presencia de feromonas en el medio extracelular* (a). *En* b, *esquema general del recurso que hacen las células de los módulos de quinasas para responder ante estímulos extracelulares y adaptarse a ellos. Estos módulos de quinasas se activan por los propios receptores, por otras quinasas o por proteínas G*

que responden de manera específica a determinadas señales extracelulares.

La primera cascada de MAP quinasas que dio en identificarse fue la vía de respuesta a feromonas en la levadura *Saccharomices cerevisiae*. Esta vía es necesaria para que las células haploides del tipo **a** puedan reconocer células del tipo α, fusionarse y formar una célula diploide (**a**/α). Cada célula haploide segrega al medio una feromona que interacciona con un receptor específico de la célula del tipo contrario.

La interacción de la feromona con el receptor, que se encuentra en la membrana celular, provoca la activación de una quinasa MAPKK(STE11) a través de su interacción con la subunidad βγ de una proteína G (asociada al receptor), mediante otra quinasa denominada STE20 o por ambos procesos.

STE11, una vez activada, fosforila una MAPKK (STE7), que a su vez fosforila y activa un par de quinasas MAP(FUS3). En este caso, la proteína STE5 mantiene asociadas las quinasas pertenecientes a los tres niveles de la cascada, mediante su unión con todas ellas. La activación de FUS3 provoca la fosforilación de STE12, un factor de transcripción responsable de la inducción de la transcripción de un conjunto de genes necesarios para la fusión de las dos células haploides.

Además, las quinasas FUS3 controlan la división celular vía la fosforilación de un inhibidor del ciclo celular (FAR1).

Por consiguiente, la interacción entre la feromona y el receptor genera una señal que viaja desde la membrana celular hasta el núcleo a través de un módulo de quinasas, y de ese modo posibilita la fusión de las dos células haploides.

En una célula de levadura hay por lo menos cinco módulos o cascadas diferentes de MAP quinasas. Cada módulo es responsable de transmitir señales específicas, generadas por diferentes estímulos del entorno. Entre esos estímulos extracelulares cabe citar, además del inducido por las feromonas, los generados por cambios en la osmolaridad o limitaciones en el contenido de nitrógeno del medio, así como por cambios drásticos de temperatura.

Las células eucariotas superiores disponen también de módulos de quinasas para transmitir señales. Detectan cambios en las condiciones extracelulares mediante alteraciones de la membrana citoplasmática o a través de receptores específicos que conllevan la activación de módulos de MAP quinasas. Ante determinados factores de crecimiento, citoquinas y diferentes tipos de estrés —irradiación ultravioleta, cambios en la osmolaridad del medio o limitación de suero—, se produce, en respuesta, la activación de cascadas de MAP quinasas; merced a éstas, la célula puede hacer frente a las nuevas condiciones y adaptarse a ellas.

FRANCESC POSAS
Instituto Oncológico Dana-Faber
Boston

## Migraña

### *Persistencia*

La migraña es una enfermedad que muestra elevada incidencia familiar. Se intuye, por tanto, la existencia de un determinante genético. Se caracteriza por crisis repetidas y altamente invalidantes. El dolor sentido presenta una distribución hemicraneal; con él concurren síntomas y signos de una disfunción vegetativa, como palidez facial, hipersensibilidad en la percepción de estímulos sensoriales (fotofobia, sonofobia y osmofobia) y manifestaciones digestivas como náuseas, vómitos e incluso despeño diarreico. Es también la norma que el migrañoso experimente una alteración afectiva expresada por irritabilidad, apatía, ansiedad y depresión.

De duración muy variable, los ataques no se prolongan más de tres días. Un carácter muy singular de la migraña es su agravación sintomática con la actividad física y con el traqueteo de la cabeza.

El impacto social de la migraña es elevado. Afecta al 17 por ciento de las mujeres y al 6 por ciento de los varones. La dolencia puede iniciarse en cualquier edad, aunque suele aparecer de forma bastante precoz. Cerca del 85 por ciento de los migrañosos la presentan antes de los 30 años. La media de edad de comienzo se sitúa entre los 17 y 19 años.

El mecanismo a través del cual se desarrolla la crisis de migraña implica la participación del sistema



*La migraña, según un grabado de comienzos del siglo* XIX

trigémino-vascular, estructura anatomofuncional clave en la modulación del dolor de cabeza. La activación del sistema puede inducirse por mensajes procedentes del córtex cerebral (áreas orbitofrontales), lóbulo límbico o hipotálamo; los mensajes generados activan estructuras del tronco cerebral, con respuesta noradrenérgica (locus ceruleus) y también serotoninérgica (núcleo dorsal del rafe y del núcleo del V par). Esta activación neurosináptica genera una respuesta que, a través del nervio trigémino, alcanza los vasos durales y desencadena la inflamación neurógena estéril.

La frecuencia de presentación de las crisis varía de un individuo a otro e incluso en un mismo individuo a lo largo de su vida. En muchos casos, su aparición se relaciona con la incidencia de factores precipitantes; entre otros, cambios hormonales, factores dietéticos, estímulos sensoriales, tensión psíquica, ruptura de ritmos circadianos y factores atmosféricos.

Como fármacos de primera elección, para la prevención de la migraña, contamos con los betabloqueantes. Su acción sobre los receptores beta adrenérgicos muestra, con una baja incidencia de efectos secundarios, un alto nivel de eficacia. Se indican también los calcioantagonistas, heterogéneo grupo de fármacos que comparten la acción de dificultar la penetración celular de iones calcio. De un modo más específico, los derivados del grupo de la piperazina resultan eficaces en la prevención de las crisis. Se recomiendan, asimismo, la metisergida, el valproato sódico y la amitriptilina.

Una vez desencadenada la crisis, pueden prescribirse ácido acetilsalicílico o paracetamol, analgésicos; entre los antiinflamatorios no esteroideos, el iboprofeno, naproxeno y diclofenaco sódico. Se debe considerar también la magnitud de los síntomas acompañantes; con frecuencia resulta necesaria la asociación de antieméticos y de ansiolíticos.

El sumatriptan, molécula afín a la serotonina, actúa como potente vasoconstrictor, sobre todo en vasos craneales y anastomosis arteriovenosas durales. Es un fármaco muy eficaz, pues bloquea el desarrollo global de la crisis.

Feliu Titus
Servicio de Neurología.
Hospital General Universitario,
Barcelona

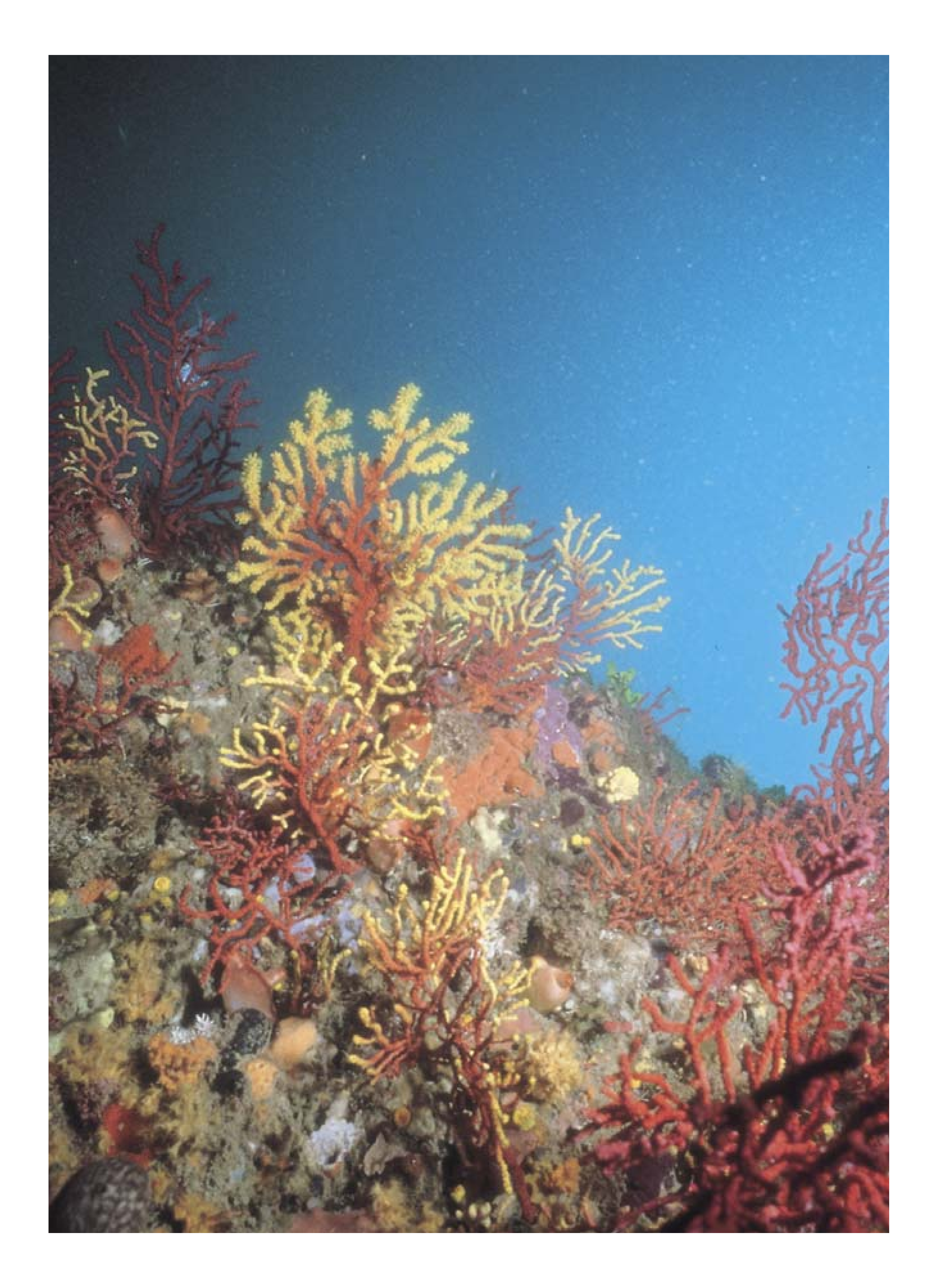
Texto y fotos: Josep-Maria Gili y Rafel Coma

## Una estructura para aprovecharlo todo

La primera impresión que recibe un buceador cuando ilumina un fondo recubierto de gorgonias en el Mediterráneo, es la de un paisaje de inusitada belleza, diversidad de formas y colorido. Si su inquietud naturalista le incita a penetrar en los detalles estructurales y en el marco ambiental donde se extienden las praderas de gorgonias, descubrirá el resultado de un proceso de evolución que ha dado lugar a una comunidad con una estructura jerarquizada muy eficaz para la captura de alimento. Diferentes tipos de lentes fotográficas (macro, 35 y 28 mm) permiten visualizar diferentes niveles de estructura tridimensional que recuerdan, en su conjunto, lo que podemos observar en un bosque mediterráneo o en una selva tropical. La comunidad de las fotografías es habitual en fondos rocosos del Mediterráneo entre 10 y 70 m de profundidad. Se desarrolla en lugares donde el régimen de corrientes locales es constante y garantiza una continuada renovación de la masa de agua inmediata sobre la comunidad. Estos organismos son suspensívoros, lo que significa que retienen todo tipo de partículas y otros organismos transportados o arrastrados por la corriente hasta la vecindad de sus cuerpos. Atrapan o filtran a través de sus cuerpos las partículas suspendidas en la masa de agua. La distribución espacial de las distintas colonias, dispuestas en dirección perpendicular a la corriente, modifica el régimen hidrodinámico (disminuye la velocidad e incrementa la turbulencia) favoreciendo la captura de partículas. La forma ramificada de las colonias en un solo plano, como la gorgonia *Paramuricea clavata* (la especie de las fotografías), hace que las colonias se comporten como filtros naturales que incrementan la superficie de captura. La disposición de los pólipos alrededor de todas las ramificaciones incrementa también las posibilidades de captura de presas. Los pólipos disponen de una corona de tentáculos ligeramente ramificados que se comportan como excelentes redes. En conjunto, se ha desarrollado una estructura tridimensional sumamente eficaz en la captura de todo tipo de presas suspendidas en la masa de agua, lo que hace que se trate de las comunidades más complejas y multiformes de los mares templados.







# Cuidado de la prole entre los

***¿Por qué en algunos insectos los progenitores arriesgan la vida por atender a sus hijos?***

**Douglas W. Tallamy**

**Fotografías de Ken Preston-Mafhan**



**1. ATENCION A LA PROLE EN LA PLUVISELVA BRASILEÑA. Se aprecia un grupo de hormigas que vigilan, solícitas y potentes, crías de homópteros de la especie *Aetalion reticulatum*. Las progenitoras han pasado a cuidarse de la puesta reciente. Las hormigas se alimentan de una secreción anal dulce, producida por las ninfas; por eso las defenderán de los posibles depredadores. Liberados de esa carga, los homópteros adultos cuidarán exclusivamente de sus huevos abandonándolos a la atención de las hormigas cuando emerjan las ninfas.**

# insectos



En el sudeste de los Estados Unidos los tíngidos del género *Gargaphia* viven en ortigas. La hembra se encarga de vigilar la puesta; de los huevos eclosionados salen las ninfas. Corren éstas un gravísimo riesgo si merodean chinches de la familia Nabidae, pues devorarán hasta la última ninfa. Los tíngidos no tiene muchas armas con que defenderse. Por eso, la madre distraerá al depredador extendiendo las alas y subiéndose a su dorso.

Mientras tanto, las ninfas corren por el nervio central de la hoja, a modo de vía rápida, para ascender, tallo arriba, por una hoja nueva, doblada todavía, donde se ocultarán. Si la hembra logra escapar, irá tras la prole y vigilará el peciolo de la hoja. Allí podrá salir al paso del depredador, que probablemente la habrá seguido. Algunas veces la hembra burlará momentáneamente al agresor; y entonces huye para guiar a las ninfas hacia la hoja más apropiada, bloqueando cualquier rama que pudieran tomar de manera errónea. Pero lo habitual es que la hembra muera en el ataque y su inmolación dé tiempo a que las ninfas escapen con vida.

**2. OCULTANDOSE en los repliegues de una hoja joven (*derecha*) las ninfas de las chinches del género *Gargaphia* se guarecen de los depredadores. Se trata de un género común del sudeste norteamericano. En otro episodio (*abajo*), la hembra se encara, sin éxito, con una larva letal de un neuróptero que está comiéndose las ninfas.**



**3. ALERTA para defender a la prole, esta hembra de chinche asesina (*Pisilus tipuliformis*) vela por sus ninfas recién emergidas (*derecha*). Chinches del género *Cocoteris* (*centro*) de Nueva Guinea y del género *Antiteuchus* de Brasil (*fondo derecha*) están a punto de perder la partida contra el depredador, pues las crías no logran ajustarse bajo el escudo protector.**



Adolph Modeer se hizo ya eco en 1764 de la atención a la prole entre los insectos. Observó que la hembra de la chinche de escudo europeo (*Elasmucha grisea*) permanecía firme sobre la puesta, abalanzando el cuerpo hacia los depredadores en vez de remontar el vuelo. Pese a ello, todavía en 1971 abundaban quienes rechazaban de plano que hubiera insectos cuidadores de la prole. Los propios que admitían la veracidad de las observaciones daban por supuesto que la protección constituía una innovación reservada a las chinches más evolucionadas.

Nos hallamos ante un comportamiento similar al atribuido a las aves y los mamíferos, es decir, formas de vida “avanzadas”. Pero la atención a la prole no es ninguna adquisición reciente. Está extendida entre los invertebrados, incluidos moluscos, gusanos, rotíferos y medusas. Por lo que se refiere a los artrópodos, este tipo de cuidados constituye la norma entre ciempiés, arañas, escorpiones, arañas de mar y crustáceos, grupo éste muy próximo al de los insectos. De hecho, la relativa rareza de insectos “paternales” —dispersos al azar a través de los 13 órdenes que existen— parece reflejar una pérdida generalizada de ese carácter que poseerían los linajes originarios.

Muchos entomólogos se preguntan por qué ha persistido este tipo de desvelo si conlleva tamaño coste ecológico. La estrategia más fácil, la seguida por la mayoría de los insectos, consiste en

**4. AGRUPADAS por seguridad debajo del cuerpo de su progenitora, las larvas de un escarabajo casidino (*Acromis sparsa*) crean una figura anular simétrica (*izquierda*). En el extremo del cuerpo portan ganchos anales en los cuales depositan sus heces consiguiendo así repeler a los depredadores incautos mediante un mordisco. La hembra se ocupa de su progenie desde el estado de huevo (*arriba*) y se encarga de pastorear a los individuos emergidos hacia las fuentes de recursos, cuidando siempre apriscarlos a todos.**





producir una puesta numerosa. En su aclamada síntesis, *Sociobiología*, Edward O. Wilson considera el cuidado una respuesta hacia un entorno insólitamente favorable o ante un medio desmesuradamente inhóspito. Sostiene allí que esta conducta debe dominar cuando abundan los recursos, y se produce una intensa competencia, cuando escasean o resulta difícil procesar el alimento, cuando las condiciones físicas son adversas y cuando la presión de depredación es asfixiante.

Doquiera que miremos, los escarabajos carroñeros y los escarabajos peloteros se defienden frente a la competición por recursos muy nutritivos, aunque efímeros, tales como la carroña y el estiércol, desarrollando una forma específica de cuidado de los hijos. Ya sean los dos, ya sea la hembra sola, ponen de inmediato el alimento a resguardo en una cavidad subterránea; de ese modo, lo protegen de sus competidores y evitan su desecación. Una pareja de escarabajos carroñeros del género *Nicrophorus* podría, por ejemplo, enterrar el cadáver de un ratoncillo y moldearlo luego en forma de copa, que alojará y alimentará a la prole. Cuando llega el momento de la eclosión de las larvas, la hembra y a veces el macho suplementan su dieta con líquidos regurgitados. Michelle P. Scott y Gonzalo Halffter han observado que los machos de estas especies impiden que otros machos les roben el botín y maten a su progenie.

En otros casos, los progenitores procesan el alimento de las crías. Las chinches del género *Sehirus* aportan semillas a las ninfas, ocultas en depresiones del suelo. Los homópteros del género *Umbonia* dejan expuestos, con una serie de rendijas espirales abiertas en la corteza, los vasos del floema de las plantas con los que alimentarán a las ninfas. Por esos tubos circulan los nutrientes. Los xilófagos deben convertir un suministro duro, indigesto y bajo en nitrógeno en una forma que las crías puedan aprovechar. Las cucarachas de la madera del género *Cryptocercus* y los escarabajos pasálidos solucionan este problema alimentando a su descendencia directamente de su ano; por ese camino obtienen fibras leñosas maceradas, protozoos (que colonizan los intestinos y metabolizan la celulosa), heces y fluidos intestinales, materiales todos ellos ricos en nitrógeno. Por otra parte, los escarabajos de la corteza la mastican y abren túneles donde dejar la puesta y donde inocularán las astillitas levantadas con hongos simbiontes que convierten la celulosa en formas digeribles para las larvas.

Los insectos cuidadores se ocupan por lo común de la puesta. En algunas especies, sin embargo, la atención se extenderá a los nuevos individuos emergidos. En este caso, debe haber una intensa intercomunicación entre padres e hijos, así como una perfecta coordinación de movimientos. Las chinches del género *Gargaphia*, las avispas portasierra y los escarabajos casidinos y escafídidos protegen a las larvas y les buscan alimento. La hembra podrá velar por la descendencia si ésta permanece agrupada; por eso las apriscará y cerrará el paso a las erráticas.

En la mayoría de los casos, las hembras se encargan del cuidado. Pero hay ocasiones en que los machos toman las riendas para que la prole pueda desenvolverse en medios adversos. Las chinches de agua, por ejemplo, ponen unos huevos de gran tamaño, que se desecarían si permanecieran sobre la superficie del agua o



**5. MADRE FORMIDABLE, la mantis religiosa *Oxyophthalmellus somalicus* (*izquierda*) yace en la base de una ramita en un desierto de Kenya. Desde esta atalaya saldrá al paso de los depredadores interesados en sus ninfas. La mantis religiosa del género *Galepsus* (*abajo*), procedente también de Kenya, ha ocultado el saco que contiene sus huevos, con lo cual logra pasar inadvertida en medio de las superficies de corteza que la rodean. De igual forma, ella permanece en guardia por si falla su artimaña.**



**6. CHINCHE PENTATOMIDA (*Tectocoris diophthalmus*) procedente de Australia. Defiende a sus huevos usando métodos agresivos. La hembra sólo produce una única puesta, con lo cual dispone de una sola oportunidad para asegurar su reproducción.**

se ahogarían si cayeran dentro. Los huevos deben, pues, mantenerse húmedos y aireados.

Es *Lethocerus* un género primitivo de chinches de agua gigantes. La hembra deposita los huevos sobre una ramita que esté encima del agua. El macho se sumerge y sale repetidas veces para remojar los huevos con las gotas que le caen y así mantenerlos húmedos. Sin olvidar que le compete también ahuyentar a los depredadores. El macho de *Belostoma*, otro género de chinches de agua gigantes que a menudo se pueden encontrar en las piscinas, transporta los huevos sobre su propio dorso; allí se los ha pegados la hembra. El macho debe mantenerlos sobre la superficie y exponerlos al aire. Moverá sus patas traseras hacia detrás y hacia delante, o se asirá a una ramita impulsándose hacia arriba durante horas, para que el agua circundante de los huevos persista aireada. De manera similar, los escarabajos estafilínidos del género *Bledius*, los carábidos del género *Bembidion* y los helódidos del género *Heterocerus* impiden que la corriente arrastre los huevos ocluyendo las aberturas de los sacos de incubación de las crías cuando sube la marea y liberándolas cuando las aguas retroceden.

Wilson ha identificado, sin resquicio para la duda, qué condiciones promueven el comportamiento paternal en los insectos. Mas, ¿por qué unos resuelven ocuparse de la prole y otros —incluso de especies emparentadas— optan, en idéntica condición, por desplegar una estrategia diferente? Podemos abordar la cuestión mediante un análisis elemental de la razón entre coste y beneficio.

Machos y hembras pagan un oneroso tributo por no remontar el vuelo y enfrentarse a los depredadores. Los riesgos son difíciles de cuantificar y apenas hay datos disponibles. Cuando hice un recuento de las posibilidades que las chinches del género *Gargaphia* tenían de sobrevivir a la depredación por arañas saltadoras, observé que las hembras ocupadas en proteger a sus ninfas presentan tres veces menos oportunidades de sobrevivir que las hembras exentas de esa responsabilidad.

DOUGLAS W. TALLAMY, experto en análisis de la interacción química entre insectos y plantas, investiga la evolución de la atención paterna en los artrópodos. Da clases en la Universidad de Delaware. La mayoría de sus trabajos se centran en las chinches del género *Gargaphia*.

El cuidado de la progenie resulta también caro porque —salvo en raras ocasiones— ata al nido a los progenitores. La producción de los huevos comporta una enorme inversión energética. Además, la hembra que monta guardia sobre la puesta no puede salir en busca de los nutrientes requeridos para una nueva tanda de huevos. Este cobro en fecundidad puede ser importante. Se ha demostrado, con hembras de *Gargaphia* retiradas experimentalmente de la protección, que éstas ponían el doble de huevos que las hembras vigilantes de la prole.

Semejante coste ha instado la adopción de estrategias alternativas, incluso dentro de la misma especie. Ciertas chinches del género *Gargaphia* y homópteros del género *Polyglypta* se ahorran riesgos y pérdidas de los cuidados colocando sus huevos sobre la puesta de otra hembra, de su misma especie si se tercia. Quedan entonces libres para una segunda puesta casi inmediata, mientras que los receptores habrán de esperar a la eclosión (en *Polyglypta*) o hasta la madurez de las ninfas (en *Gargaphia*). Si una hembra de *Gargaphia* no logra colocar sus huevos en otro nido, reducirá los riesgos si reserva la defensa agresiva de la prole para cuando ya es mayor y tenga poco que perder o bien para cuando las ninfas se acercan a los últimos estadios de desarrollo y es probable que lleguen a la madurez.

Los homópteros del género *Pubilia* limitan los costes de la atención de un modo peculiar, observado por Catherine M. Bristow. La hembra permanece con su prole hasta que las hormigas descubren el grupo y empiezan a consumir una secreción anal azucarada producida por las ninfas. En ese momento, la madre se independiza y cede el cuidado a las hormigas, muy capacitadas para defender las ninfas ante depredadores.

Si costosísimo es para la hembra el cuidado de la prole, para el macho resulta prohibitivo. Si nos fijamos en la fisiología, la producción de esperma resulta barata. Por tanto, aunque la atención a los hijos le robe tiempo para la búsqueda de alimento, la reducción de nutrientes no debería comportar una drástica merma en la capacidad fecundadora. La gabela a pagar aquí es la promiscuidad; mientras permanezca ligado al nido, no tendrá libertad para ir en busca de otras hembras a las que fecundar.

A esa cortapisa hemos de agregar otro inconveniente, el de la imposibilidad de garantizar su paternidad a la



**7. EL MACHO DE LA CHINCHE de agua gigante (*Abedus herberti*) habitante del cañón de Sycamore, en Arizona, transporta con él los huevos que ha fecundado. La hembra se los ha dejado pegados al dorso. El macho —una de las pocas especies en las que él se ocupa de la atención a los hijos— va y viene para mantener a los huevos húmedos y aireados.**

mayoría de los machos. Las hembras retienen almacenado el esperma y pueden escoger el de qué macho quedar fecundadas. Tal inseguridad sobre la transmisión de los propios caracteres evidencia el dudoso negocio de la atención para la mayoría de los machos.

No debe, pues, sorprendernos la extrema rareza de un cuidado reservado en exclusiva a los machos. Se da sólo en tres familias de heterópteros. Los machos se las ingenian para librarse de esa contribución en algunas especies de la familia Reduviidae (chinches asesinas), en un número incluso menor de chinches de la familia Coreidae y en todas las chinches de agua gigantes (familia Belostomatidae).

*Rhinocoris*, chinche de la familia Reduviidae, hace ostentación de su cuidado del nido. Ante ese despliegue, diríase que las hembras merodeantes dan por supuesto que el macho vigilante de su prole se concentra en esa misión y, por tanto, lo buscan para la cópula. Puesto que las hembras rehusarán aparearse con machos que no estén cuidando a su progenie (excepto en la etapa temprana de la estación, cuando decae la puesta), los machos pugnarán por masas de huevos que proteger. Tal comportamiento resulta rentable, pues las hembras hacen la puesta luego de la cópula, si no durante la misma; ello significa que el macho acepta ocuparse de los que con más seguridad son progenie suya.

En ambos tipos de chinches, de agua y asesinas, es muy alta la densidad de hembras que buscan macho. Lisa Thomas ha estudiado la especie *Rhinocoris tristis*, chinche asesina procedente de Kenya, en la que el cuidado se deja al macho. El insecto vive en plantas del género *Stylosanthes*; bebe el néctar de sus flores y caza a sus presas que se mueven por los alrededores. Puesto que las chinches se concentran en una misma planta y las hembras pueden encontrar sin dificultad un macho para el cuidado, éste no ve coartada su promiscuidad. Robert L. Smith ha descrito observaciones similares en las poblaciones, relativamente densas, de chinches de agua gigantes confinadas en los estanques: el macho no necesita ir en busca de hembra; muchas saldrán a su encuentro.

Pese a tan extenso muestrario, la mayoría de los insectos evitan los costes de la vigilancia y optan por recurrir a mecanismos que aseguren la supervivencia de los huevos. Ovipositores punzantes (apéndices ensiformes usados para hacer la puesta) o duras e impenetrables cubiertas para los huevos permiten ocultar la progenie en tejidos vegetales de plantas o bien encerrar a sus huevos en grietas o sellarlos en oquedades naturales al abrigo de posibles daños. En el trasfondo de todas estas innovaciones hallamos el desarrollo de la reproducción iterópara. En vez de poner todos los huevos a la vez —en un acto de reproducción semélpara— para ocuparse de ellos luego, los insectos han optado en su mayoría por puestas pequeñas, distanciadas en el tiempo y el espacio.

Por sí sola, la estrategia es ya una vía muy eficaz para ahorrar pérdidas. Si un depredador descubre una puesta, sólo podrá acceder a una fracción del total de los huevos producidos por la hembra. Si las chinches del género *Gargaphia* no protegen sus puestas —superiores al centenar de huevos— los depredadores acabarán con un 56 % antes de la eclosión. Por contra, *Corythucha ciliata*, chinche de la familia Tingidae, pone los huevos en 33 ocasiones y los distribuye en diferentes hojas de su planta hospedadora. Merced a su estrategia, pierde únicamente el 16 % de sus huevos arrebatados por los depredadores.

¿Por qué, pues, no han abandonado todos los insectos el cuidado vigilante de la prole? Volvamos al análisis de la razón de coste a beneficio. El progenitor paga un precio importante por cuidar a sus crías, sólo si por ellos se resiente la fecundidad. No hay tal coste cuando la crudeza del invierno o la escasez de los recursos frenan la producción de huevos. En tal caso el cuidado constituye una opción óptima.

Las hembras de *Parastrachia japonensis*, chinche excavadora, alimentan a las crías de frutos caídos de los árboles del género *Schoepfia*. Deben confinar su época reproductora al momento de máxima abundancia del fruto. En ese breve intervalo cuenta la hembra con alimento suficiente para producir una puesta prolífica, que protegerá juntamente con las provisiones necesarias para poder subsistir durante varias semanas, sin tener que sacrificar subsiguientes oportunidades reproductoras.

Casi todas las especies que cuidan de la prole se ven obligadas a una sola puesta debido a los cambios estacionales, escasez de recursos u otras restricciones de tipo ecológico. Considerada la atención a la prole enmarcada en la reproducción iterópara y la semélpara, resulta más fácil explicar la ausencia de una pauta, que tanto desconcertó durante años a los entomólogos. Para la mayoría de los insectos, la posibilidad de ampliar su período reproductor en el tiempo y en el espacio ha hecho que el cuidado sea prohibitivo e innecesario. Para quienes tienen pocas posibilidades de cruzarse, ésa podría ser la forma de asegurar que su descendencia logre llegar a la madurez y siga transmitiendo sus genes.

BIBLIOGRAFIA COMPLEMENTARIA


COSTS AND BENEFITS OF THE EGG-DUMPING ALTERNATIVE IN *GARGAPHIA* LACE BUGS (*HEMIPTERA: TINGIDAE*). D. W. Tallamy y L. A. Horton, en *Animal Behavior*, vol. 39, páginas 352-360, febrero 1990.

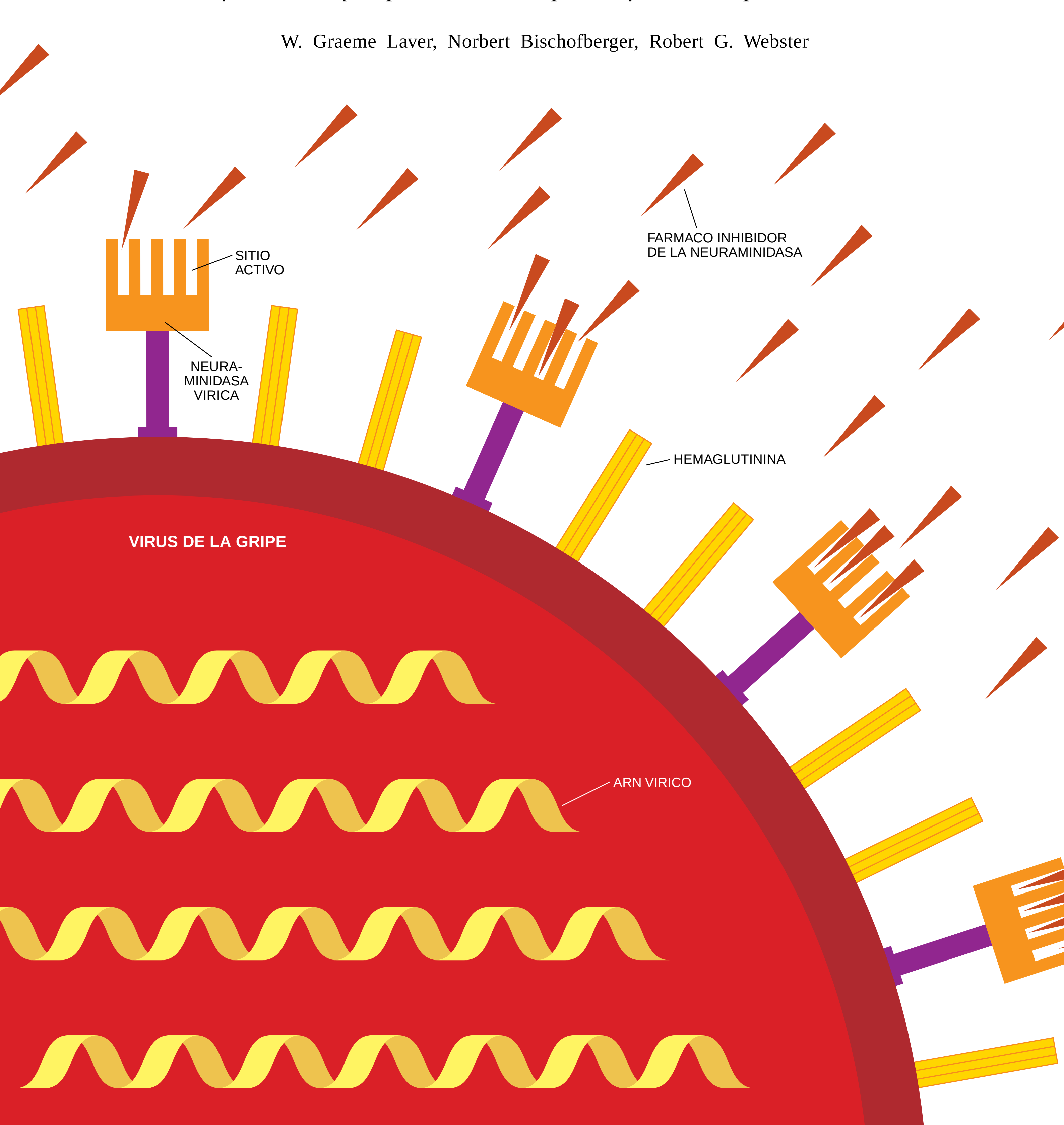
THE EVOLUTION OF PARENTAL CARE. T. H. Clutton-Brock. Princeton University Press, 1991.

THE EVOLUTION OF SOCIAL BEHAVIOR IN INSECTS AND ARACHNIDS. Dirigido por Jae C. Choe y Bernard J. Crespi. Cambridge University Press, 1997.

# Desarme de los virus de la gripe

*Pronto contaremos con medicinas que tratarán la gripe mediante la detención de la replicación vírica en los tejidos, fármacos que podrían desempeñar funciones preventivas*

W. Graeme Laver, Norbert Bischofberger, Robert G. Webster



No es raro que una cepa insólita de virus de la gripe comience a pasar, de repente, de un sujeto a otro. Por su propio carácter insólito, son contadas, si alguna, las personas que han desarrollado inmunidad adquirida en exposiciones anteriores; hasta los vacunados carecen de defensas efectivas. Las vacunas contra la gripe protegen de variantes del virus que los expertos suponen que serán activas en una estación dada, pero no de otras variantes inesperadas. Al no encontrar nada que le ponga freno, la nueva cepa se disemina intacta, sembrando enfermedad y muerte por todo el planeta.

La peor pandemia (epidemia mundial) que se recuerda es la de 1918. Mató a más de veinte millones de personas, en algunos casos pocas horas después de manifestarse los primeros síntomas. Al desastre, desencadenado por el virus de la "gripe española", siguieron en 1957 una epidemia de "gripe asiática", la gripe de Hong Kong en 1968 y la "gripe rusa" en 1977. (Los nombres reflejan la impresión popular sobre los lugares donde se inició la pandemia, aunque ahora se sospeche que los cuatro episodios y, posiblemente, la mayoría de los demás se originaron en China.)

Los expertos en salud pública advierten de que en cualquier momento puede aparecer otra pandemia y que ésta podría ser tan peligrosa como el episodio de 1918. En 1997 se detectaron en Hong Kong 18 casos de una variante letal de virus de la gripe que terminó con la vida de seis de ellos, en lo que se temió fuera el principio de la siguiente ola. Las autoridades regionales consiguieron controlar el problema con rapidez. Detectaron la fuente de contagio —pollos, patos y gansos infectados— y sacrificaron toda la población aviar de Hong Kong.

La próxima vez puede que la humanidad no tenga tanta suerte. Si una cepa tan letal como la de Hong Kong se diseminara por nuestras comunidades densamente pobladas, un 30 % de la humanidad podría morir —por acción directa del virus o por infecciones bacterianas secundarias— antes de que diera tiempo a desarrollar una vacuna con la que proteger a los que inicialmente se hubieran salvado. Se necesitan unos seis meses para desarrollar una vacuna contra una variante cualquiera de virus de la gripe, probar su seguridad y distribuirla, demasiado tiempo para una pandemia que se extiende con tanta rapidez.

Si la temida pandemia tarda aún otro año o incluso más en aparecer, puede que ya se disponga de nuevos métodos para evitar tanta enfermedad y muerte. A finales de este año y después de superar extensos ensayos clínicos, se podría aprobar la comercialización de dos fármacos como nuevas armas contra la gripe. Los agentes zanamivir (Relenza) y GS 4104 son prometedores en la prevención de la gripe y en la reducción de la duración y severidad de los síntomas en personas que empiezan el tratamiento cuando ya están enfermas.

A diferencia de las vacunas, que estimulan el sistema inmunitario para que impida que los virus ganen pie en el organismo, y a diferencia de los remedios caseros habituales, que suavizan los síntomas pero que no tienen efecto sobre la infección, estos fármacos han sido diseñados para atacar directamente al virus. Mediante el bloqueo de la acción de la neuraminidasa, una enzima vírica decisiva, reducen notablemente la proliferación del virus en el organismo. Otros inhibidores de neuraminidasa se encuentran también en estudio, aunque no se ha llegado a la fase de ensayo clínico.

Ya existen en el mercado dos fármacos, amantadina y rimantadina, con acción contra los virus de la gripe. Pero estos dos agentes, cuyo modo de acción se sustenta en mecanismos distintos, presentan graves inconvenientes. Pueden causar confusión y otros efectos secundarios neurológicos y no son eficaces contra una de las dos clases principales de virus, el tipo B, que afectan a los humanos. Aún más, los virus de la gripe parecen desarrollar sin mayor dificultad resistencia contra estos fármacos. Así, los individuos que en las primeras fases de una epidemia han recibido este tratamiento pueden transmitir versiones del virus resistentes a la medicación. Este problema es especialmente acuciante en las comunidades cerradas, como las casas de convalecencia y residencias de ancianos.

La historia de cómo se desarrollaron estos nuevos fármacos es una maravillosa combinación de suerte y lógica. El desciframiento en 1983 de la estructura tridimensional de la neuraminidasa es el punto de partida más inmediato. Pero fue una serie anterior de descubrimientos la que permitió que los científicos identificaran una parte específica de la molécula de neuraminidasa como un talón de Aquiles en todas las variantes de virus de la gripe, una debilidad que fármacos bien construidos podrían explotar.

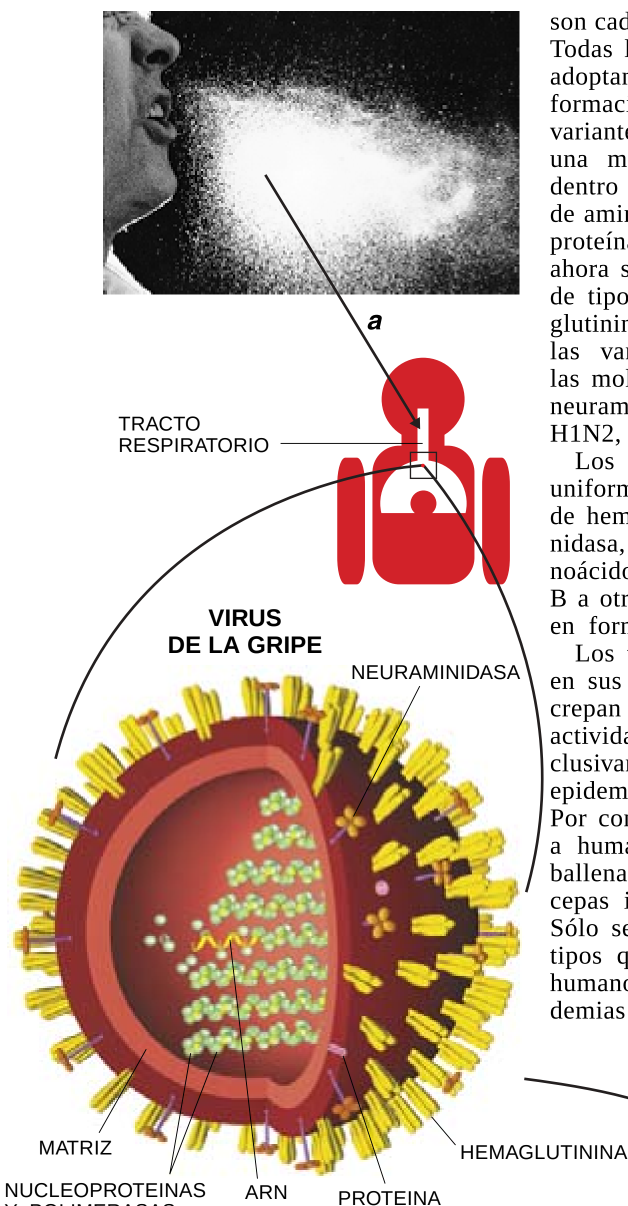
Una de esas líneas de investigación puso al descubierto propiedades esenciales de los virus de la gripe y de sus estrategias de pervivencia. Saben los biólogos que los virus son, a grande rasgos, genes envueltos en proteínas que protegen esos mismos genes o que participan en la replicación vírica en el organismo. A veces, como en el caso del virus de la gripe, estos constituyentes están también ceñidos por una membrana lipídica. Las enfermedades causadas por virus suelen comenzar con la invasión de tipos celulares seleccionados, en cuyo interior el virus se replica antes de salir al exterior para infectar otras células. Los síntomas se deben a la alteración de la función de las células colonizadas, secundaria a la proliferación vírica, y a los esfuerzos del sistema inmunitario por contener la infección, que producen inflamación, fiebre y dolores sistémicos.

Las cepas que colonizan humanos revelan especial afinidad por las células epiteliales que tapizan el tracto respiratorio. Si la infección tiene éxito, en uno o en un par de días aparecerán los síntomas clásicos de nariz tapada, mucosidad abundante, tos seca, estornudos, fiebre, dolores, cansancio agudo y pérdida del apetito. Las descripciones históricas basadas en síntomas indican que las epidemias de gripe han azotado a la humanidad desde antes del siglo V a.C.

En 1933 se aisló la primera cepa de virus de la gripe a partir de un humano. Desde entonces, se sabe que los virus gripales pertenecen a dos tipos básicos —A y B— que se diferencian en ciertas proteínas internas. No parece que un tercer tipo (C) cause enfermedad grave.

Los virólogos clasifican las variantes del tipo A según dos proteínas de superficie que sobresalen en espícula, la hemaglutinina y la neuraminidasa (enzima diana de los nuevos fármacos). En cuanto proteínas

**1. CON LOS NUEVOS FARMACOS (*puntas, en color rojo*) se avanzará en la prevención y tratamiento de la gripe. Se unen al sitio activo catalítico de la neuraminidasa, una molécula que sobresale de la superficie de los virus gripales. La enzima capacita a las partículas víricas recién formadas el tránsito de una célula a otra. Si se inactiva la enzima, el virus queda inmovilizado.**




son cadenas de aminoácidos plegadas. Todas las variantes de hemaglutinina adoptan esencialmente la misma conformación tridimensional y todas las variantes de neuraminidasa presentan una morfología característica. Pero dentro de cada grupo, la secuencia de aminoácidos constituyentes de una proteína puede variar mucho. Hasta ahora se han identificado en el virus de tipo A quince subtipos de hemaglutinina y nueve de neuraminidasa; las variantes se denominan según las moléculas de hemaglutinina y de neuraminidasa que presenten: H1N1, H1N2, H2N2, etc.

Los virus de tipo B, mucho más uniformes, portan una sola variante de hemaglutinina y otra de neuraminidasa, si bien la secuencia de aminoácidos también varía de una cepa B a otra. Cada subtipo A se presenta en formas levemente diferentes.

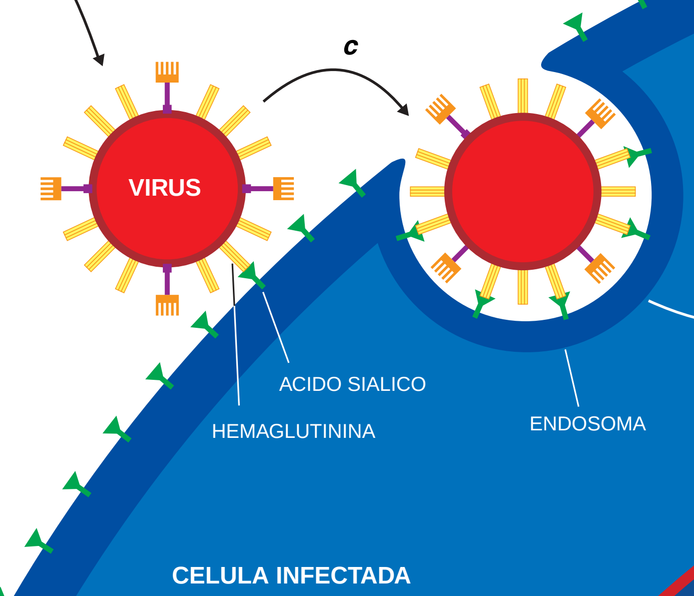
Los tipos A y B no sólo divergen en sus características químicas. Discrepan también en el alcance de su actividad. El tipo B, que infecta exclusivamente a humanos, suele causar epidemias regionales, no pandemias. Por contra, la gripe de tipo A afecta a humanos, cerdos, caballos, focas, ballenas y aves, aunque no todas las cepas infectan a todas las especies. Sólo se han identificado cuatro subtipos que provoquen enfermedad en humanos: los responsables de las pandemias sufridas en nuestro siglo.

A pesar de las diferencias, ambos tipos de virus desarrollan el mismo ciclo biológico. Para que una copia o partícula de virus entre en la célula, la hemaglutinina vírica debe enlazarse con el ácido siálico, molécula glicosada de la superficie celular. Una vez introducido, el virus queda inicialmente secuestrado en una especie de burbuja. Pronto se liberan sus genes, constituidos por cadenas de ARN, y las proteínas internas, listos para llegar al núcleo celular.

Allí, algunas proteínas víricas comienzan a replicar las cadenas de ARN vírico y a sintetizar ARN mensajero, una forma de ácido nucleico que lee y traduce a proteínas la maquinaria celular encargada de la síntesis de proteínas. Las nuevas moléculas sintetizadas, proteínas y ácidos nucleicos, se unen y crean una partícula vírica nueva antes de salir al exterior.

Desafortunadamente para el virus, las partículas emergentes permanecen recubiertas por ácido siálico, sustancia de la superficie celular en que se anclan los virus gripales. Si el ácido siálico persistiera en el virus y en la superficie de la célula destinada a la replicación vírica, la hemaglutinina de las partículas recién formadas se enlazaría al ácido siálico, provocando que las partículas víricas se agolparan sobre la célula, como insectos pegados en el atrapamoscas. No podrían alcanzar nuevas células.

Pero el virus guarda un as en la manga. La moléculas de neuraminidasa de las partículas víricas recién formadas seccionan las moléculas de ácido siálico. En otras palabras, las




**2. CICLO BIOLOGICO del virus de la gripe. A menudo implica la transmisión de gotitas, emitidas en un estornudo, de las vías aéreas de una persona a otra (*a*). Una partícula vírica (*detalle, arriba*) penetra en una de las células que tapizan el tracto respiratorio después de que una molécula de hemaglutinina del virus se una al ácido siálico de la superficie celular (*b*). Esta unión induce la entrada intracelular del virus (*c*) y la llegada del material genético vírico, compuesto por ARN, y de sus proteínas internas al núcleo celular (*d* y *e*). Algunas de estas proteínas colaboran en la replicación del ARN (*f*) y en la producción de ARN mensajero que la maquinaria celular encargada de la síntesis de proteínas utiliza como molde para fabricar las proteínas víricas (*g* y *h*). Luego, las proteínas y los genes del virus se ensamblan para formar nuevas copias del virus o partículas víricas (*i*) que salen de la célula. Las partículas emergen recubiertas de ácido siálico. La presencia de esta sustancia en la superficie del virus y de la célula determinaría que las moléculas de hemaglutinina de una partícula quedaran pegadas al ácido siálico de partículas vecinas y de la célula; los nuevos virus permanecerían pegados a la célula. Pero la neuraminidasa del virus rompe el ácido siálico de las superficies (*j*) y libera a las nuevas partículas (*k*), que invadirán otras células.**

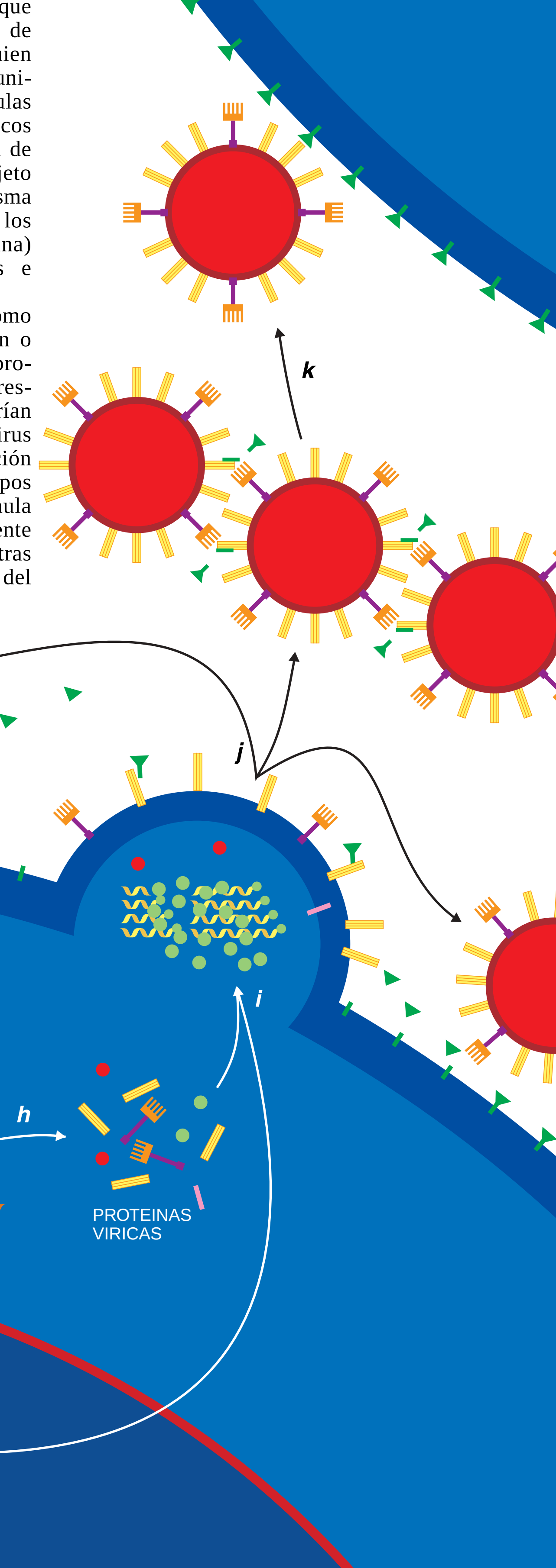
espículas de neuraminidasa disuelven el “pegamento” de ácido siálico y dejan libertad de movimientos a las partículas. La enzima facilita también el desplazamiento del virus por el moco que hay entre las células de las vías aéreas.

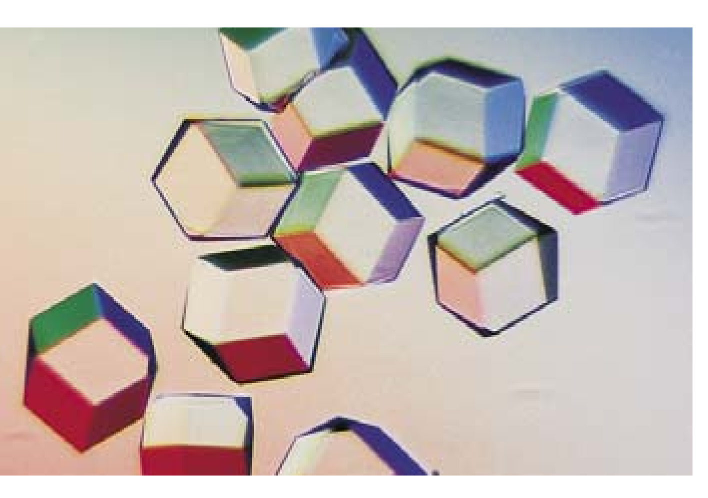
Hacia 1960 se sabía ya que el fármaco que bloqueara cualquier etapa del proceso de replicación impediría que el virus causara enfermedad o permitiría controlar una infección activa. Pero se desconocía cómo llevar la idea a la práctica. Conscientes de que el virus de la gripe crecía en el interior de las células cuya maquinaria empleaba para fabricar sus proteínas víricas, los biólogos temían que los agentes destructores del virus afectaran de paso a las células sanas.

Continuó la investigación sobre las razones por las que algunas cepas producían epidemias locales y otras desataban terribles pandemias. Se llegó a la determinación de que, para ser útil contra la gripe, el fármaco debía hallarse capacitado para atacar a todas las variantes del virus, incluidas las que se ignora si provocan enfermedad en humanos.

Una cepa de virus de la gripe puede desendenar una epidemia, local o global, sólo en el caso de que las personas expuestas carezcan de inmunidad adecuada. Cuando alguien contrae la gripe, su sistema inmunitario produce anticuerpos, moléculas que reconocen segmentos específicos de hemaglutinina y neuraminidasa de la superficie del virus. Si ese sujeto sufre una nueva exposición a la misma cepa, los anticuerpos (sobre todo los dirigidos contra la hemaglutinina) bloquearán prestamente al virus e impedirán que siga infectando.

Si el virus no cambiara nunca, como sucede con el virus del sarampión o el de las paperas, los anticuerpos producidos durante la infección o en respuesta a una vacuna proporcionarían inmunidad permanente. Pero los virus de la gripe están en transformación incesante. Por tanto, los anticuerpos producidos un año resultan de nula o escasa eficacia si en la siguiente estación de gripe se enfrentan a otras variantes del virus. La cuantía del





**3. CRISTALES DE NEURAMINIDASA obtenidos de una variante del virus de la gripe que afecta a las aves salvajes de la Gran Barrera de Arrecifes de Australia. Gracias a ellos, se ha determinado la estructura tridimensional de la enzima y diseñado fármacos que se unen a su sitio activo. Los colores de la imagen son reflejo de la luz coloreada.**

cambio determina, en buena medida, si una epidemia queda controlada o se propaga sin freno.

Los virus de la gripe se van modificando a través de la ruta de la "deriva" antigénica, una revisión gradual de la secuencia de aminoácidos de una proteína (antígeno) capaz de desencadenar una respuesta inmunitaria. Estas alteraciones aparecen por pequeñas mutaciones en el gen de la proteína. Unas veces, la mutación comporta cambios sutiles en la estabilidad o en la actividad de una proteína. Otras, atenta contra la proteína y dificulta la viabilidad del virus. Pero puede también reforzar la pervivencia, mediante la reconfiguración de un sitio de la hemaglutinina que hasta entonces era reconocido por un anticuerpo.

W. GRAEME LAVER, NORBERT BISCHOFBERGER y ROBERT G. WEBSTER han realizado notables aportaciones al control de las epidemias y pandemias de gripe. Laver, profesor de la Universidad Nacional Australiana de Canberra, aisló los primeros cristales de neuraminidasa. Bischofberger es vicepresidente de investigación de Gilead Sciences en Foster City. Webster fue, junto con Laver, el primero en rastrear las cepas de virus de la gripe en animales inferiores. Desempeña la cátedra Rose Marie Thomas y dirige el departamento de virología y biología molecular del Hospital de Investigación Pediátrica St. Jude de Memphis.

Cuando los genes y las proteínas de hemaglutinina o neuraminidasa han acumulado alteración tras alteración, se tornan punto menos que irreconocibles para la mayoría de los anticuerpos. Se desencadena así una nueva epidemia, cuya extensión se detiene cuando encuentra grupos humanos cuyo sistema inmunitario ya se había enfrentado a algunas de tales alteraciones.

Según parece, el tipo B no conoce otra vía de modificación que la deriva antigénica. Evoluciona gradualmente en su huésped humano para ocultar su naturaleza al repertorio inmunitario de una población. Las cepas del tipo A pueden sufrir, por contra, cambios más drásticos. Esas variaciones mayores se compendian en el "desplazamiento" antigénico, capaz por sí solo de desatar una pandemia.

Cuando se presenta un desplazamiento antigénico, las cepas desarrollan espículas de hemaglutinina nuevas; a veces, con una molécula incluida de neuraminidasa con la que la mayoría de las personas no han tenido nunca contacto. De ese modo, el virus elude el repertorio antigénico entero de todas las poblaciones del mundo y desencadena una pandemia. Con la facilidad de transporte intercontinental, un viajero puede diseminar un virus peligroso e inédito, de una parte a otra de la tierra, en un solo día.

Un cambio tan drástico no puede resultar de una simple mutación genética. El proceso mejor estudiado de desplazamiento antigénico se basa en la coincidencia de dos cepas víricas en una célula huésped. De esa guisa, los genes empaquetados en las nuevas partículas de virus (y sus correspondientes proteínas) provienen en parte de una cepa y en parte de otra. Cabe semejante reordenación genética porque el genoma de los virus de la gripe consta de ocho cadenas discretas de ARN (cada una de ellas determina una o dos proteínas). Estas cadenas se pueden mezclar y complementar cuando se forman nuevas partículas de virus A en células dualmente infectadas. Hay, en efecto, virus de la gripe que infectan a personas y a cerdos; si un cerdo se viera infectado por un virus humano y una cepa que infecta sólo a aves, el cerdo podría terminar por producir una cepa híbrida, igual en todo al virus humano menos en presentar una molécula de hemaglutinina propia del virus aviar.

Pero el desplazamiento antigénico acontece, asimismo, cuando un virus gripal de animales, incapaz hasta entonces de producir infección en humanos, de pronto se vuelve virulento para las personas.

Nadie sabe qué clase de desplazamiento antigénico desencadenó la pandemia de gripe española de 1918, causada por el subtipo H1N1 de virus de la gripe A. Se ha comprobado que la reordenación genómica fue responsable de las pandemias de la gripe asiática de 1957 y de la gripe de Hong Kong de 1968, causadas respectivamente por H2N2 y H3N2. Algunos estudios sugieren que las aves acuáticas pudieron aportar genes poco frecuentes y que los cerdos sirvieron probablemente de banco de mezcla. Este papel de los cerdos explicaría por qué las pandemias acostumbran brotar en China: millones de aves, cerdos y personas viven en estrecha relación.

H5N1, el virus que mató seis personas en Hong Kong en 1997, no había emergido, sin embargo, de ninguna reordenación genómica. Pasó directamente de las aves al hombre. Pero H5N1 no pudo transmitirse de un humano a otro. Si se le hubiera dado tiempo suficiente para adquirir, por mutación o reordenación, esa capacidad, quizás habría escapado en seguida a todo control.

El toque de alerta de 1997 ha convencido a los expertos en salud pública de que no basta con seguir el curso de la gripe en humanos, cosa que ya se viene haciendo. Hay que vigilar también los casos que se dan en aves migratorias. Es probable que estas aves sirvan de reservorio de cepas del tipo A de año en año, que luego se diseminan a las aves domésticas y a otras especies. La rápida identificación de cepas animales con potencial de dañar a las personas podría ayudar a evitar un desastre de salud pública.

El incidente de Hong Kong ha avivado el interés por estudiar la naturaleza de la "barrera de las especies", que impide que muchas cepas de la gripe puedan pasar de un animal a otro. Si se conociera mejor ese obstáculo, podrían sellarse los huecos por donde se cuelan cepas animales para infectar al hombre.

Del estudio sobre biología de la gripe acometido a principios de los ochenta se desprendía que el fármaco ideal debería bloquear la actividad de alguna molécula relacionada con el ciclo reproductor del virus mediante la ocupación de un sitio "conservado" de la molécula efectora. La medicina tendría que hacer blanco en una región cuya secuencia de aminoácidos persistiera constante en todas las cepas. Actuaría así contra cualquier virus de la gripe que apareciera en personas, incluidos los transmitidos directamente por animales.

El trabajo estructural que permitió desarrollar los inhibidores de la neuraminidasa surgió a raíz de un descubrimiento accidental. A finales de los setenta, uno de nosotros (Laver) se aprestaba a determinar si la espícula N2 del virus que había causado la pandemia gripal de Hong Kong de 1968 (H3N2) procedía de la cepa responsable de la pandemia de gripe asiática de 1957 (H2N2); se proponía, pues, comparar la secuencia de aminoácidos de las moléculas. Aisló y concentró las cabezas, esto es, los dominios que emergen de la superficie del virus.

Cuando Laver liberó de los virus purificados las cabezas de neuraminidasa y las concentró en una centrífuga, halló que el amasijo de material resultante no era amorfo como suele suceder con las proteínas. Observó cristales, distribuciones ordenadas de moléculas, esenciales para descifrar la estructura tridimensional de proteínas grandes. En concreto, la inesperada producción de cristales de neuraminidasa implicaba que se podría descifrar la estructura de la neuraminidasa.

Eso mismo es lo que el grupo de Peter Colman, de la Organización de Investigación Científica e Industrial Commonwealth (CSIRO) de Australia, realizó en 1983. El trabajo reveló que las espículas de neuraminidasa de los virus de la gripe constaban de 4 monómeros, moléculas idénticas entre sí. El tetrámero resultante se asemejaba a cuatro globos colocados encima de un eje. El eje, único, penetraba la membrana vírica y los globos sobresalían. El grupo de Colman no tardó en descubrir que cada monómero de neuraminidasa mostraba una hendidura central profunda en su superficie.

Se observó que, si bien las moléculas de neuraminidasa podían diferir en la secuencia de aminoácido concreta, todas las versiones conocidas, tanto del tipo A como del B, presentaban una similitud impresionante. Los aminoácidos que tapizaban la pared de la hendidura eran invariantes.

Cuando alguna parte de una molécula se resiste al cambio, el motivo habitual es que se trata de componentes esenciales para la función de la molécula. En este caso, la uniformidad apuntaba a que la hendidura formaba el sitio activo de la neuraminidasa que rompía el ácido siálico y que los aminoácidos conservados de la hendidura eran imprescindibles para ejecutar la acción catalítica. El trabajo posterior confirmó estas hipótesis.

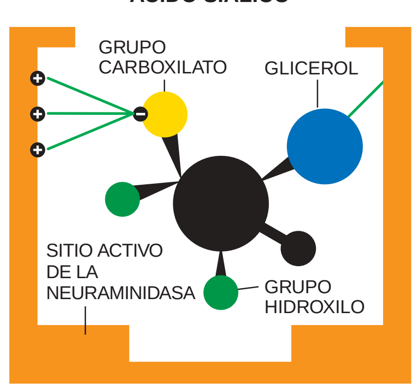
Dado que los virus de la gripe no pueden diseminarse de una célula a otra sin la acción de la neuraminidasa, los nuevos descubrimientos implicaban que un fármaco capaz de ocupar y bloquear el sitio activo podría inhibir la neuraminidasa en todas las versiones del virus de la gripe. Esto es, la "llave" serviría de cura universal para la gripe.

Persiguiendo idea tan sugestiva, el grupo de Colman identificó los aminoácidos del sitio activo que establecen contacto con el ácido siálico. También determinó qué aminoácidos de la hendidura no se unían al ácido siálico, pero podían servir de anclaje al fármaco. Observó, por ejemplo, que la hendidura tenía tres aminoácidos con carga positiva que se enlazaba con firmeza a un grupo carboxilato del ácido siálico con carga negativa.

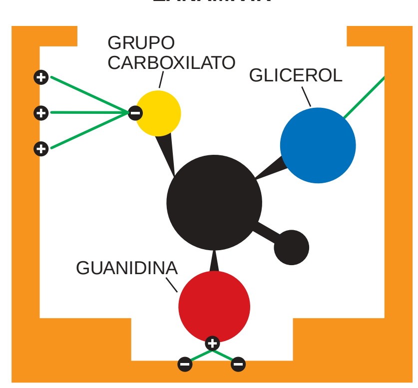
Había, además, en el fondo de la hendidura una bolsa donde se alojaban dos aminoácidos de glutamato con carga negativa que no entraban en contacto con el ácido siálico, aunque estaban presentes en todas las neuraminidasas examinadas. Un grupo hidroxilo (OH) del ácido siálico apuntaba hacia esa bolsa suplementaria, sin llegar a alcanzarla.

Estas características sugerían que, si se cambiaba el grupo OH por un grupo atómico grande, con carga positiva, se podría obtener un derivado que se uniera firmemente. El grupo positivo se acomodaría quizás en la bolsa suplementaria de la base del sitio activo; se anclaría en los grupos glutamato de la bolsa dotada de

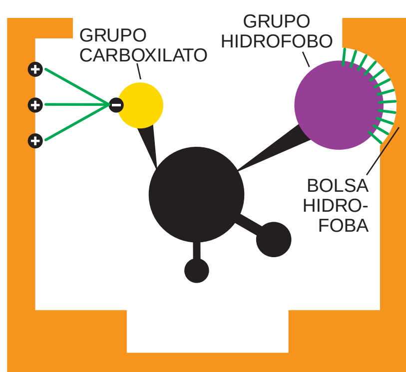
ACIDO SIALICO




ZANAMIVIR




GS 4104 Y GS 4071




**4. LOS NUEVOS FARMACOS, al unirse al sitio activo de la neuraminidasa, lo inactivan. Se enlazan mejor que el propio ácido siálico, la molécula que dicha enzima degrada. El ácido siálico (*izquierda*) se mantiene trabado al sitio activo, una hendidura, mediante sus grupos glicerol y carboxilato, que establecen enlaces (*líneas verdes*) con los aminoácidos del sitio activo. El zanamivir (*centro*) presenta engarces adicionales mediante la sustitución de grupo hidroxilo de un derivado del ácido siálico por un grupo guanidina, de gran tamaño y con carga positiva, que a su vez establece fuertes enlaces con los dos aminoácidos con carga negativa del fondo de la hendidura. El GS 4104 se convierte en GS 4701 en el interior del organismo. La molécula resultante (*derecha*) conserva las uniones del grupo carboxilato al igual que el ácido siálico, pero utiliza también un grupo hidrófobo. Este grupo induce la formación de una bolsa hidrófoba similar en la hendidura, bolsa que mantiene fijo el fármaco mediante atracciones hidrófobas (*rayitas verdes*).**

## El perfeccionamiento de las vacunas

Ante el futuro de los fármacos antigripales, cabría preguntarse si no se controlaría mejor la enfermedad con una vacuna universal, capaz de evitar la infección mediante la inducción de una respuesta inmunitaria protectora del organismo contra cualquier cepa. Para nuestro infortunio, no existe todavía nada parecido.

Los inmunólogos buscan la manera de acelerar la fabricación de vacunas, de modo que sea posible una inmunización rápida en cuanto aparezca un brote virulento. Se trabaja también en vacunas no inyectables en su administración, para facilitar su aceptación y estimular la inmunización de los niños, principales responsables de la diseminación de la gripe.

Las vacunas antigripales se han generalizado desde 1940. El proceso de manufactura comienza con el análisis de muestras del virus recogidas en 110 sitios de vigilancia repartidos por la tierra. La Organización Mundial de la Salud propone en febrero tres cepas, dos del tipo A y una del B, responsables presumibles de la mayoría de los casos de gripe de la estación siguiente (de noviembre a marzo, en el hemisferio Norte). Estas serán la base de la vacuna.

Para elaborar una vacuna bastaría en principio con dejar crecer un gran número de cepas seleccionadas, inactivarlas para que no puedan producir infección y combinarlas en una preparación única. Pero las cepas seleccionadas suelen crecer lentamente en el laboratorio y resulta difícil producir grandes cantidades. Para superar este obstáculo, se insertan proteínas que estimulan el sistema inmunitario: hemaglutinina y neuraminidasa de la superficie de cepas seleccionadas en una variante de virus de la gripe que crezca rápida en el laboratorio. Se infectan embriones de pollo con dos virus, uno de los seleccionados y el de crecimiento raudo. El proceso se itera en cada cepa seleccionada. En poco tiempo, en el embrión crecen muchas partículas víricas que presentan las espículas de hemaglutinina y neuraminidasa correspondientes a las cepas que se supone causarán la epidemia de ese año. Se aíslan esos reagrupamientos de intenso crecimiento y se entregan a los fabricantes de vacunas, que las producen en masa en embriones de pollo.

Tiempo atrás, las variantes inactivadas de estos virus, e incluso las reordenaciones de las tres cepas seleccionadas, servían de vacuna. La mayoría de los laboratorios llevan ahora el proceso un paso más lejos. Rompen el virus y fabrican la vacuna directamente con las proteínas víricas. Las proteínas inducen inmunidad pero son incapaces de producir infección. Cualquier tipo de estas vacunas estimula la formación de anticuerpos, que se unen a los virus infecciosos portadores de tales proteínas e intervienen en su eliminación.

Un enfoque alternativo se basa en la preparación de vacunas a partir de virus debilitados. Los virus estimulan la síntesis de anticuerpos y la producción de linfocitos *T*, leucocitos que reconocen y eliminan células infectadas por

carga negativa y sin emplear hasta entonces.

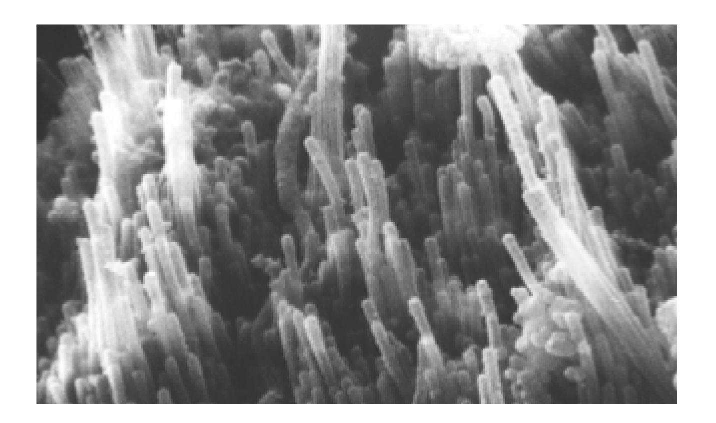
Después de varios intentos, en 1993 Mark von Itzstein y su equipo, de la Universidad Monash de Melbourne, hallaron que, si se cambiaba el grupo OH del ácido siálico por un grupo guanidina, grande y dotado de carga positiva, se obtenía un inhibidor de las neuraminidasas gripales potentísimo. El inhibidor no ejercía apenas efecto sobre enzimas afines sintetizadas por bacterias y mamíferos, señal de que el compuesto no lesionaría las células humanas.

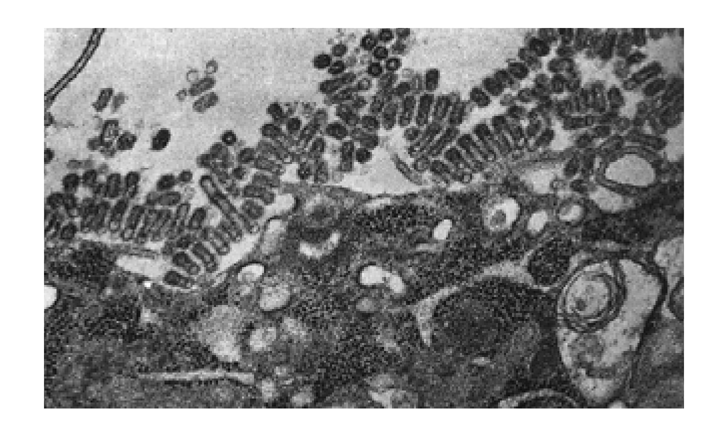
Los estudios en animales y los ensayos preliminares en humanos revelaron que la sustancia, el zanamivir, prevenía los síntomas gripales en individuos que luego resultaron infectados por el virus y que remitía la gravedad de los síntomas en quienes tomaron el preparado después de infectarse. Este no se administra en pastilla, sino inhalado por la nariz o la boca hacia el sistema respiratorio.

El grupo guanidina, gracias al cual el zanamivir constituye un óptimo inhibidor, impide que se tome en pastilla. Los fármacos ingeridos han de cruzar las células que tapizan el intestino y alcanzar el torrente sanguíneo, para distribuirse por otras partes del cuerpo. Pero a las moléculas dotadas de carga les cuesta atravesar las membranas celulares, lipídicas y permeables sobre todo a sustancias sin carga.

La inhalación es la forma habitual de administración de medicinas que han de operar en el tracto respiratorio. Glaxo Wellcome, de Stebenage, inició más pruebas en humanos con zanamivir. Algunos pacientes, sin embargo, preferían la pastilla. Por eso, Gilead Sciences, de Foster City, se alió con F. Hofman-La Roche, de Basilea, para fabricar un inhibidor de neuraminidasa que se administrara por vía oral. Además, si el fármaco circula por la sangre podría combatir cualquier virus de la gripe virulento que infectara células fuera del tracto respiratorio.

Tras sintetizar y someter a ensayo varios compuestos, Gileat y F. Hoffman-La Roche se concentraron en GS 4071, tan potente como zanamivir. Del análisis estructural se desprendía







virus. Las células *T* reaccionan ante cepas que estén emparentadas, a diferencia de los anticuerpos, que sólo reconocen una sola cepa. Las vacunas de virus vivo proporcionarían inmunidad duradera mientras las variaciones de una cepa del virus de la gripe se limitaran a pequeños cambios en la estructura de sus moléculas de superficie.

**LA VACUNACION ANUAL constituye la mejor forma de evitar la gripe. Llegará un día en que se podrán sustituir las inyecciones por vaporizadores nasales.**

La empresa Aviron, de Mountain View, ha creado una vacuna de virus vivos administrada en forma de vaporizador nasal. Tras haber superado las pruebas en humanos, niños incluidos, saldrá muy pronto al mercado. Ocurre, sin embargo, que su fabricación no es directa como la de virus muertos, ni aportaría probablemente una defensa rápida en una situación de pandemia súbita.

Para reducir el tiempo de producción, se ensayan métodos de fabricación que evitan la fecundación obligada de grandes cantidades de huevos. Se pueden insertar los genes de hemaglutinina y neuraminidasa de cepas de virus de la gripe seleccionadas en otro tipo de virus, como los baculovirus, de rápido crecimiento en células de cultivo, algo en que el virus de la gripe se muestra díscolo. Cuando los virus manipulados genéticamente se replican en las células, fabrican cantidades notables de las proteínas de la cepa seleccionada, que luego pueden purificarse y aprovecharse como vacunas. Se tarda dos o tres meses en preparar y distribuir tales vacunas recombinantes, pero falta por establecer su eficacia.

Otra estrategia de vacunación capaz de proporcionar un producto aún más rápido utiliza ADN "desnudo". Se insieren los genes de hemaglutinina y neuraminidasa en anillos de ADN. En teoría, si se inyectaran dichos plásmidos en la piel o el músculo, entrarían en las células de la vecindad, que los emplearían para fabricar proteínas del virus de la gripe. Luego, presentarían estas proteínas en su superficie para que las detectara el sistema inmunitario. En respuesta fabricaría anticuerpos y células T capaces de neutralizar virus libres y erradicar cualquier célula infectada. Las vacunas antigripales de ADN desnudo han dado buen resultado en animales de laboratorio, pero aún no se han ensayado en personas.

*—W.G.L., N.B. y R.G.W.*

que tenía también un grupo principal de unión a la neuraminidasa que se encontraba en el zanamivir: un grupo carboxilato dotado de carga negativa que se unía a los aminoácidos con carga positiva del sitio activo de la enzima. Además, la interacción entre GS 4071 y el sitio activo inducía la rotación de un aminoácido de la hendidura y creaba una bolsa hidrofóbica nueva. Esta bolsa repelente del agua servía de anclaje para un componente hidrofóbico similar (una cadena de átomos de carbono e hidrógeno) del fármaco.

La sustancia funcionó bien en el tubo de ensayo. No dañaba las células sanas. Pero fracasó en una prueba crucial: el grupo carboxilato negativo le impedía atravesar el intestino de los animales y llegar a la sangre. Mas bastó un pequeño reajuste —se recubrió el grupo negativo— para solucionar el problema. La forma con envoltura GS 4104 arribaba sin problemas a la circulación y se liberaba de su envoltura en la sangre y el hígado. Convertido de nuevo en su forma original (GS 4071) inhibía la actividad de la neuraminidasa y vedaba la diseminación del virus en el tracto respiratorio de los animales de ensayo. Lo mismo que el zanamivir, dio muestras de funcionar bien en humanos.

**5. LAS NUEVAS partículas víricas emergen de la célula en filamentos (*izquierda*), aunque también pueden ser esféricas o presentar cualquier forma intermedia. Cuando se inactivan sus moléculas de neuraminidasa, se observa la acumulación de partículas en la superficie de la célula (*pequeños husos y esferas, arriba en la micrografía de la derecha*).**

Los informes del último otoño sobre los exhaustivos ensayos clínicos controlados en torno a zanamivir y GS 4104 confirmaron y extendieron los descubrimientos iniciales. Si se inicia el tratamiento con zanamivir o la forma ingerida GS 4104, día y medio después del inicio de los síntomas, el tiempo que la gente se encuentra enferma se reduce en un 30 % (uno y medio a tres días). Los compuestos también reducen la gravedad de los síntomas. Un ensayo con GS 4104, por ejemplo, indicó que los diarios de valoración de los pacientes mostraban que los síntomas eran entre un 25 % y un 60 % más suaves que en los pacientes que tomaron placebo. Además, los fármacos disminuyen el riesgo de infecciones bacterianas secundarias —así, la bronquitis—, potencialmente letales, en un 50 % o más. Estas complicaciones son una causa muy importante de muertes por gripe, sobre todo entre los ancianos y las personas con enfermedades concomitantes.

Cierto estudio restringido de zanamivir manifiesta que la forma inyectable del fármaco también es eficaz. Son a su vez prometedoras las investigaciones en el terreno de la prevención. Cuando se probó el zanamivir, el 6 % de los que no lo tomaron, por sólo el 2 % de los que lo tomaron, contrajeron la gripe. GS 4104 produce resultados similares. Hasta la fecha ninguno de los dos fármacos ha producido efectos secundarios graves.

A partir de estos estudios Glaxo Wellcome ha solicitado autorización para comercializar la forma inhalada de zanamivir en Australia, Europa, Canadá y los EE.UU. Se espera que Gilead y F. Hoffman-La Roche hagan lo mismo con el GS 4104 en EE.UU. y Europa a lo largo del año en curso. Otros inhibidores de la neuraminidasa desarrollados por BioCryst Pharmaceuticals de Birmingham han funcionado en estudios con animales y se ha autorizado a Johnson & Johnson para un posterior desarrollo.

Para el observador ocasional, el recortar pocos días la duración de la gripe puede parecerle un triunfo menor. Pero se trata de algo más profundo de lo que sugiere a simple vista. La fatiga y otras incomodidades que se sufren durante los últimos días de la gripe resultan, sobre todo,

## ¿Qué determina la virulencia?

Cuando la gripe mata, acostumbran caer personas con un sistema inmunitario afectado: ancianos y enfermos. Pero en ocasiones siega la vida de personas jóvenes y vigorosas con la misma rapidez y contundencia con que acaba con los débiles.

Sucedió en la pandemia de gripe española de 1918. Causó más bajas que las producidas en el frente durante la Primera Guerra Mundial. Mientras perduró la pandemia, hasta los soldados más fornidos perecieron. Algunos se empezaron a encontrar mal por la mañana, pasaron la tarde en cama y murieron por la noche. También eran jóvenes y vigorosos las víctimas de 1997 y, a pesar de ello, murieron seis de las dieciocho personas infectadas por una cepa de la gripe en Hong Kong.

¿Qué convierte a una cepa en más letal que otra? Parte de la respuesta reside en la capacidad de infectar tejidos diferentes de los habituales, que son el tracto respiratorio de los mamíferos y el gastrointestinal de las aves. Los microbiólogos interesados por la transformación virulenta buscan características determinantes de que ciertas cepas sean promiscuas y se tornen pantrópicas respecto a los tipos celulares que atacan.

Hace ahora quince años, Webster y su equipo del Hospital de Investigación Pediátrica St. Jude de Memphis dieron con una pista. En 1983, en Pennsylvania, un virus que había provocado una enfermedad gastrointestinal leve en pollos, empezó de repente a arrasar granjas enteras. La investigación demostró que todo obedecía a una simple sustitución de un aminoácido por otro en la hemaglutinina, proteína de la superficie vírica. Este cambio mínimo había permitido al virus replicarse y dañar órganos de las aves afectadas.

El trabajo subsiguiente reveló que este mínimo cambio estructural había tenido un profundo efecto en la actividad vírica. Cuando el virus de la gripe entra en una célula, queda secuestrado en una burbuja intracelular (endosoma). Se reproduce porque las moléculas de hemaglutinina de la superficie vírica facilitan que la membrana vírica se fusione con el endosoma. Cuando esto sucede, los genes y las proteínas del virus salen del endosoma e inician la producción en masa de nuevas copias del virus. La hemaglutinina facilita la fusión sólo si se parte en dos antes de que el virus entre en la célula. De ello son responsables

**1. LA PANDEMIA DE GRIPE ESPAÑOLA que comenzó en 1918 mató a más de 20 millones de personas y condenó a los familiares de las víctimas a hacer largas colas para recibir alimentos. La escena ocurrió en Cincinnati.**



de la actividad residual del sistema inmunitario, tras la eliminación del organismo de la mayoría de los virus. Los fármacos antigripales no pueden afectar esta actividad, pero al obstruir la replicación vírica los compuestos sí pueden recortar la duración y aligerar la gravedad de la parte inicial, la más desagradable de un episodio de gripe.

Cuanto antes se empiece el tratamiento, mejores resultados se cosecharán, en parte porque el organismo albergará una cantidad de virus menor y más domeñable y porque los fármacos no pueden curar las lesiones que el virus inflige a los tejidos antes de que se inicie la terapia. De la combinación de tales observaciones con la investigación desarrollada en el campo de la prevención se infiere que la mejor forma de usar tales fármacos puede ser en conjunción con pruebas instantáneas que detecten la presencia de una infección incipiente antes de que la persona desarrolle síntomas. No es disparatado pensar que, en un futuro cercano, las personas podrían aplicarse, cada mañana, una tirita de papel de prueba de detección del virus de la gripe en la lengua. Si la tira cambia de color, se sabrá que nuestro organismo aloja ya al virus de la gripe y que hay que comenzar a tomar un fármaco antigripal para controlar los síntomas. Existen protocolos de detección precoz en el consultorio médico y se está avanzando en adaptarlo a su empleo domiciliario.

Estos fármacos funcionan sólo ante virus de gripe. Importa, pues, determinar si una persona ha contraído la gripe y no un simple resfriado u otra enfermedad, para evitar el desperdicio de emplearlas en combatir resfriados, causados por otros tipos de virus, alergias o infecciones bacterianas que producen síntomas pseudogripales.

Aunque los agentes bloqueantes de neuraminidasa han generado enorme expectación, quedan cuestiones pendientes. Por lógica, la inhibición de la replicación vírica y la reducción de las complicaciones secundarias deberían salvar vidas, pero no sabemos todavía si el fármaco realmente evita muertes.

Aún más, las drogas que los humanos empleamos contra virus y bacterias acostumbran terminar superadas por las resistencias. En su lucha por la supervivencia, estos microorganismos desarrollan variantes en las que se altera la diana del fármaco, para hacerla irreconocible y poder escapar a los efectos del fármaco. La amantadina y la rimantadina son algunos ejemplos de fármacos contra los que se han desarrollado resistencias. ¿Existen razones para pensar que los nuevos fármacos antigripales no terminen igual?

Las hay. Se ha intentado producir en el laboratorio cepas de virus de la gripe resistentes al zanamivir y al GS 4104. Hasta ahora sólo han tenido un éxito limitado. Algunas cepas ganaron resistencia, como se predecía, alterando los aminoácidos del sitio activo de la neuraminidasa, pero estos cambios hicieron tambalear la estabilidad de la enzima o su actividad normal, lo que nos adelanta las dificultades que rodearían a la pervivencia del virus en el organismo.

ciertas enzimas de la familia de las serinproteasas que se sintetizan en el tracto digestivo de las aves y en el respiratorio de los mamíferos, pero que son menos frecuentes en otros tejidos.

La sustitución del aminoácido encontrada en ese virus letal para las aves alteraba el sitio de fragmentación de la hemaglutinina; lo hacía accesible a otras enzimas que también escinden proteínas (proteasas furinomorfas) comunes en los tejidos. El incremento de la susceptibilidad a la partición permitía la infección sistémica por el virus.

El descubrimiento sugería que la pandemia humana de 1918 alcanzó semejante potencia destructora porque la cepa responsable presentó una forma mutante de hemaglutinina susceptible de ser partida por proteasas comunes, encontradas fuera de las vías respiratorias humanas. Para corroborarlo y buscar otras fuentes de virulencia, el grupo de Jeffery Taubenberger estudió el material genético obtenido de tres víctimas de esa epidemia: dos soldados (de los que se habían conservado muestras de tejido) y una mujer inuit, cuyo cuerpo fue exhumado del permafrost de Alaska en agosto de 1997.

**2. EN FEBRERO DE 1998 se tuvo que transportar a Hong Kong millones de pollos, tras haber sacrificado la producción propia para evitar la diseminación de un virus de la gripe letal. En la imagen, los funcionarios se disponen a tomar muestras de sangre de las aves.**



El trabajo genético reveló la secuencia de aminoácidos de la molécula de hemaglutinina. El grupo de Taubenberger ha secuenciado otra molécula de superficie, la neuraminidasa. La molécula de hemaglutinina parece ser anodina en lo que se refiere al sitio de partición. Además, Taubenberger afirma que sus análisis del gen de la neuraminidasa indican que tampoco en esta molécula se observa otra mutación que se había propuesto como posible fuente de promiscuidad vírica. Tan interesante hipótesis proponía que una cierta mutación de la neuraminidasa podría permitirle almacenar serinproteasas y utilizarlas para partir la molécula de hemaglutinina en tejidos carentes de tales proteasas.

Hasta ahora, y tal vez para siempre, las razones de la extrema virulencia de la pandemia de 1918 siguen siendo un misterio. Pero los investigadores tienen una idea de por qué la cepa de la gripe que apareció en Hong Kong en 1997 fue tan letal.

El virus originado en las gallinas portaba una variante de hemaglutinina muy sensible a la partición. Pese a todo, no hay pruebas inequívocas de que esta característica fuera la razón de la virulencia. Se buscan otras explicaciones plausibles.

*—V.G.L., N.B. y R.G.W.*

---

Otras cepas no basaron su estrategia de resistencia en modificaciones de la neuraminidasa, sino en cambios químicos en la hemaglutinina. Recuérdese que el virus de la gripe necesita la neuraminidasa para romper el ácido siálico de la superficie de nuevas partículas víricas al objeto de que la hemaglutinina no quede pegada al ácido siálico de las partículas vecinas y se evite así la diseminación hacia otras células. El cambio en la hemaglutinina redujo su afinidad por el ácido siálico, obviando la necesidad de que interviniera la neuraminidasa. Aunque apareció en células de cultivo, el segundo grupo de mutantes no mostró resistencia a los fármacos en animales. Cabe imaginar que el enlace maltrecho de la hemaglutinina y el ácido siálico redujo la infectividad de las cepas al debilitar su capacidad de anclarse en las células.

La investigación no se detiene en los fármacos bloqueantes de la neuraminidasa. Desde hace tiempo se sueña, sin éxito, con fármacos cuya diana fuera el sitio de unión de la hemaglutinina con el ácido siálico. La eficacia de la amantadina y de la rimantadina se descubrió antes de conocer su mecanismo de acción. Ahora se sabe que interfieren la actividad de M2, proteína vírica que funciona como un canal iónico. La inhibición de M2 explica por qué los fármacos no ejercen efecto sobre el tipo B de la gripe: estos virus carecen de moléculas M2. A la par que se avanza en la investigación de nuevos fármacos, se intenta mejorar las vacunas contra la gripe.

Entre una pandemia y otra menudean epidemias circunscritas importantes. En 1994, según el Centro de Control y Prevención de Enfermedades de los EE.UU., se estimó que 90 millones de norteamericanos, alrededor del 35 % de la población, contrajo la gripe. En conjunto, estas personas pasaron 170 millones de días en cama y perdieron 69,3 millones de días de trabajo. Sólo en ese país la gripe afecta de un 10 a un 20 % de la población cada año y causa alrededor de 20.000 muertes por complicaciones relacionadas con la gripe.

Estas cifras, como las asociadas con las pandemias, es probable que disminuyan en los próximos años cuando se disponga de los nuevos fármacos antigripales y de vacunas que tengan una mayor difusión y se fabriquen con mayor rapidez. La sociedad puede acercarse a una era en que la especie humana le gane la partida al temido virus de la gripe.

BIBLIOGRAFIA COMPLEMENTARIA

THE PLAGUE OF THE SPANISH LADY. Richard Collier. Atheneum, 1974.

EPIDEMIC AND PEACE, 1918. Alfred W. Crosby, Jr. Greenwood Press, 1976.

INFLUENZA. Edwin D. Kilbourne. Plenum Medical Book Company, 1987.

INFLUENZA VIRUS. Frederick G. Hayden y Peter Palese en *Clinical Virology*. Dirigido por D.D. Richman, R.J. Whitley y F.G. Hayden. Churchill Livingstone, 1997.

TEXTBOOK OF INFLUENZA. Dirigido por Karl G. Nicholson, Robert G. Webster y Alan Hay. Balckwell Science, 1998.

# Microsatélites de ADN

*Las secuencias repetitivas de ADN desempeñan un papel destacado en la adaptación de las bacterias a los ambientes donde medran. ¿Cumplen una función similar en organismos superiores?*

E. Richard Moxon y Christopher Wills

**EL MICROSATELITE (*región resaltada*) ilustrado consta de una secuencia de unidades de ADN o bases —*CAAT* (citosina-adenina-adenina-timina)— que se repite cinco veces. Cada secuencia de *CAAT* se empareja con su complementaria *GTTA* (guanina-timina-timina-adenina) en la hebra opuesta de la escalera de ADN. Las secuencias repetitivas, o motivos, pueden contener en los microsatélites hasta seis bases, y cada secuencia puede aparecer en múltiples copias.**



E. RICHARD MOXON y CHRISTOPHER WILLS abordan el estudio de los microsatélites de ADN desde enfoques complementarios. Moxon, pediatra, es experto en enfermedades infecciosas; Wills, biólogo, se dedica al estudio de la evolución.

El código genético humano está constituido por unos 3000 millones de bases de ADN. Sólo de un 10 a un 15 por ciento de esas bases forman parte de los genes, los planos maestros que la célula utiliza para construir sus proteínas. A otras secuencias de pares de bases les competen —en el hombre y otros muchos organismos— funciones muy importantes; por ejemplo, promover la activación o inactivación de los genes y mantener unidos los cromosomas. Queda una buena parte del ADN sin manifiesta misión obvia. Algunos la llaman "chatarra".

En ese "ADN chatarra" hay secuencias de características singulares, agrupadas bajo la denominación común de ADN satélite. Lo forman, en efecto, secuencias repetitivas de las cuatro bases del ADN —adenina (*A*), citosina (*C*), guanina (*G*) y timina (*T*)—, en diversa combinación y repetidas en una suerte de tartamudeo. Los microsatélites, que contienen las repeticiones más cortas, encierran un interés muy superior al que cabría deducir de su tamaño; realizan funciones múltiples y sorprendentes.

Se hace cada vez más patente que la naturaleza repetitiva del ADN microsatélite le capacita para crecer o disminuir en longitud y que estos cambios pueden acarrear consecuencias buenas o malas para los organismos portadores.

En determinadas bacterias patógenas las secuencias repetitivas promueven la aparición de propiedades nuevas que capacitan a esos microorganismos para sobrevivir ante cambios del entorno potencialmente letales. Es probable que algunos microsatélites ejerzan efectos importantes en el hombre, pues hay en el genoma humano —el conjunto del ADN de cada una de nuestras células— unos 100.000. Hasta la fecha, la única misión asignada a los microsatélites del hombre es negativa; causan diversas enfermedades neurológicas. Sin embargo, podrían ser restos supervivientes de procesos evolutivos que desembocaron en la conformación de nuestra especie.

Mientras unos indagan los motivos de la cuantiosa presencia de ADN repetitivo en el hombre, otros se apoyan en los microsatélites para el diagnóstico de enfermedades neurológicas y para la detección de personas en riesgo de padecerlas. Se está comprobando que los microsatélites cambian de longitud en fases precoces de ciertos cánceres, lo que les convierte en valiosos marcadores para el diagnóstico precoz de tales patologías. Y puesto que la longitud de los microsatélites varía de una persona a otra, en ellos comienzan a apoyarse la identificación de criminales y la determinación de la paternidad. A este procedimiento se le denomina "identificación dactilar por ADN" ("fingerprinting").

En los años sesenta se conoció el primer ADN satélite. Se descubrió que, cuando se centrifugaba ADN en determinadas condiciones, el ácido nucleico se disponía en dos o más capas: una banda o estrato principal, que contenía los genes, y bandas secundarias, o bandas satélites. Las bandas satélites constaban de secuencias muy largas de ADN repetitivo. En 1985 Alec J. Jeffreys encontró otras regiones más cortas de ADN repetitivo; tales minisatélites —así los llamó— eran zonas repetidas de 15 o más bases. (El grupo de Jeffreys determinó que el número de repeticiones en un minisatélite difería de un individuo a otro, y en ello basó su diseño de la técnica de la identificación por ADN.) A finales de los ochenta, James L. Weber y Paula L. May, por un lado, y Michael Litt y Jeffrey A. Lutty, por otro, aislaron satélites constituidos por zonas repetitivas de ADN menores incluso. Los denominaron microsatélites. Se han mostrado muy eficaces para la identificación de las personas por ADN.

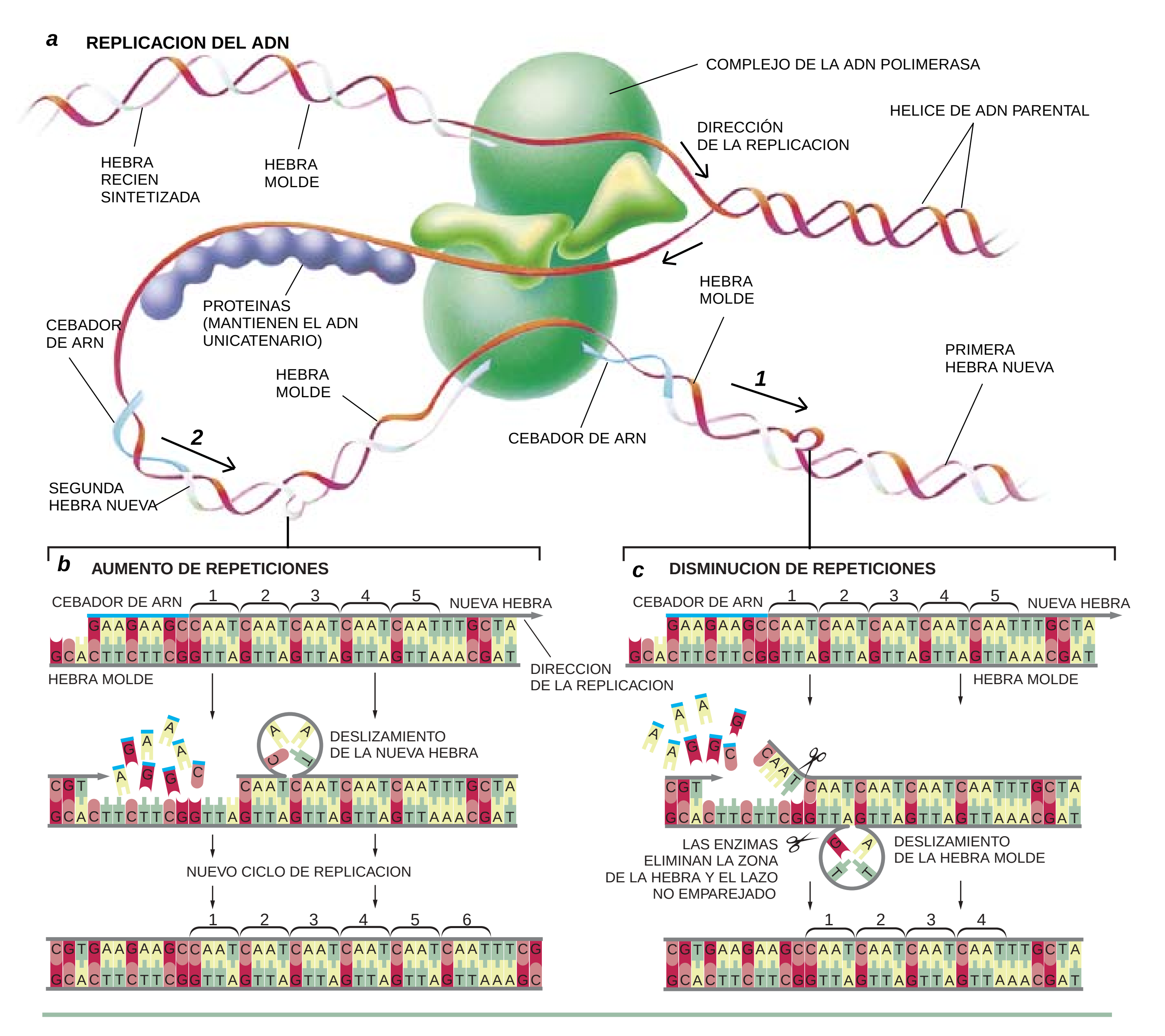
Se admite que el ADN microsatélite está constituido por secuencias de hasta unos seis pares de bases, iterados y unidos en secuencia continua. Debe su interés especial en los procesos evolutivos a una altísima tasa de mutación: la probabilidad de que un microsatélite adquiera o pierda, de una generación a la siguiente, una de las regiones repetidas multiplica 10.000 veces la probabilidad de que tal ocurra en el gen responsable de la anemia falciforme, en el que la mutación de una base da lugar a la enfermedad. Y si es raro que esa mutación singular de la anemia falciforme revierta el gen a su estado inocuo, no puede predicarse lo mismo de los microsatélites, que vuelven sin problema a sus longitudes previas, incluso a las pocas generaciones.

En 1986 el grupo de Thomas F. Meyer sacaba a la luz la función de los microsatélites en las bacterias patógenas. El equipo de Meyer investigaba *Neisseria gonorrhoeae*, la bacteria causante de la gonorrea, una enfermedad de transmisión sexual. *N. gonorrhoeae*, unicelular, posee una familia de hasta 12 proteínas de la membrana externa determinadas por genes *Opa*. (Les viene el nombre de la "opacidad" que presentan las colonias bacterianas que sintetizan proteínas Opa.) Merced a las proteínas Opa, la bacteria se adhiere e invade a las células epiteliales —las que tapizan el tracto respiratorio— y los fagocitos, células del sistema inmunitario. Cada gen *Opa* contiene un microsatélite compuesto por múltiples copias de un motivo de cinco bases, *CTCTT*.

El enorme potencial de variación que arrastran las repeticiones de los microsatélites obedece a su proclividad a equivocarse en la replicación del ADN; cometen un error frecuente: el desajuste por deslizamiento entre secuencias complementarias. Antes de que una célula —bacteriana o de otro tipo— se divida, ha de duplicarse su ADN. En este proceso complicado, cada molécula

## Errores en la formación de pares por culpa del deslizamiento

En este proceso, el número de repeticiones de microsatélites aumenta o disminuye cuando una célula copia su ADN antes de dividirse. Durante la replicación del ADN (*a*) las enzimas que constituyen el complejo de la ADN polimerasa abren la doble hélice parental de ADN y copian ambas hebras. Una de las copias se fabrica de retales: el complejo de la polimerasa sintetiza un fragmento corto (*1*) comenzando por el cebador de ARN, después salta hacia delante para generar un segundo fragmento también corto (*2*). Cuando la polimerasa termina la síntesis de este segundo fragmento, se elimina el cebador de ARN, y los dos fragmentos establecen contacto con ADN. Se produce un aumento del número de secuencias repetitivas de microsatélites (*b*) cuando la nueva hebra se desliza hacia atrás sobre una de las repeticiones en la hebra molde; ello determina que la polimerasa agregue una repetición extra en la nueva hebra para llenar el hueco. Se acorta (*c*) cuando la hebra vieja se desliza, dando lugar a que las enzimas de reparación eliminen una repetición.




de ADN es una doble hélice en escalera de caracol, cuyos peldaños corresponden a los pares de bases. El código genético se escribe con las bases de un lado de la escalera; las del otro lado son complementarias (*A* siempre se empareja con *T*, y *C* con *G*).

Durante la replicación del ADN, la escalera se parte por la mitad de cada peldaño; se separan los pares de bases al alejarse ambas cadenas helicoidales. Las ADN polimerasas copian cada una de las hebras. Mientras se va construyendo la nueva hebra, ésta se empareja con la que le sirve de molde. Pueden darse errores de replicación por corrimiento cuando la hebra vieja, el molde o la hebra que se está formando se deslizan

y establecen un emparejamiento inadecuado con la zona repetitiva de la otra hebra. Este deslizamiento es la causa de que la ADN polimerasa añada o elimine una o más copias de las zonas repetitivas de la nueva hebra de ADN.

La frecuencia de este mecanismo de deslizamiento alcanza valores altos en *N. gonorrhoeae*. En cada tanda de división bacteriana, de cada 100 o 1000 células hijas habrá una que porte una mutación que modifica el número de repeticiones de *CTCTT*. Esta modificación puede ejercer efectos drásticos sobre los genes *Opa*, porque la información genética se interpreta en "palabras" de tres bases, los codones. Las proteínas son ristras de aminoácidos y cada codón determina qué aminoácido ha de incorporarse en la cadena proteica. Como la secuencia repetitiva no está constituida por tres bases, un aumento o una disminución del número de estas secuencias desplaza el significado de los codones subsiguientes.

En el caso de los genes *Opa*, la deleción de una secuencia *CTCTT* conduce a la producción de una proteína recortada. No podrá ya adherirse a la célula hospedadora; en consecuencia, la bacteria portadora de la proteína acortada queda incapacitada para penetrar en las células. Pero el deslizamiento subsecuente abre la posibilidad de recuperar la zona repetitiva, permitiendo, por tanto, al gen *Opa* producir de nuevo una proteína funcional.

Conocido por variación de fase, este cambio reversible se da en muchas bacterias causantes de enfermedades. Con los cambios alternativos que llevan a la conexión y desconexión de los genes *Opa* de una generación a otra, *N. gonorrhoeae* incrementa sus posibilidades de supervivencia. Hay momentos en que al microorganismo le conviene adherirse a la célula hospedadora y penetrar en ella. Tal acontece cuando la bacteria se propaga y llega a un nuevo huésped. En otras ocasiones le resulta a ésta más ventajoso no interaccionar con la célula huésped; en particular si se trata de fagocitos, que destruyen las bacterias tras atraparlas en su interior.

Las implicaciones del emparejamiento incorrecto por deslizamiento y su relación con la capacidad para modificar las moléculas de superficie se han estudiado *in extenso* a propósito de la bacteria *Hemophilus influenzae*. Las células del tipo b de esta bacteria pueden provocar una meningitis infecciosa que afecta al cerebro y puede resultar letal. Hasta la introducción de la vacuna a finales de los ochenta, uno de cada 750 infantes menores de cinco años contraía la meningitis por *H. influenzae*.

La membrana externa del *H. influenzae* está festoneada con moléculas de lípidos y azúcares unidas entre sí para formar un lipopolisacárido (LPS). Una parte del LPS, el colín-fosfato, facilita la adhesión de *H. influenzae* a las células de la mucosa de la nariz y de la garganta, donde la bacteria se instala sin producir síntomas. Al menos tres de los genes requeridos para la síntesis del LPS contienen microsatélites construidos a partir de la secuencia de cuatro bases *CAAT*. Igual que sucede con los genes *Opa* de *Neisseria gonorrhoeae*, los cambios aquí en el número de repeticiones de *CAAT* pueden hacer que el *H. influenzae* sintetice LPS dotado de colín-fosfato o sin él.

Jeffrey N. Weiser ha demostrado que las cepas de *H. influenzae* que tienen colín-fosfato (ChoP) en sus moléculas de LPS, las cepas ChoP+, colonizan nariz y garganta de un modo más eficaz que las cepas que carecen de ella, las cepas ChoP–. Sin el colín-fosfato, la bacteria resiste mejor la agresión letal de factores diversos vehiculados por la sangre

## En busca de papá Chimpancé

Se ha recurrido a microsatélites de ADN para identificar criminales a través del análisis de las huellas nucleotídicas. Se emplean también esas secuencias para conocer la vida sexual de los animales, en el marco de programas conservacionistas de especies en peligro de extinción.

La huella de ADN, peculiar de cada individuo, revela que la longitud de las secuencias nucleotídicas de los microsatélites difiere en cada persona. Se obtienen esas huellas mediante enzimas especiales, que permiten fabricar millones de copias exactas de varios microsatélites de un individuo. Luego, en un gel se separan las copias en razón del tamaño. Se configura así un patrón de bandas, similar al código de barras.

Pascal Gagneux, David S. Woodruff y Christophe Boesch han empleado microsatélites de ADN como trazadores para sondear los hábitos de apareamiento de un grupo de chimpancés del bosque Taï de Costa de Marfil. Han recogido pelos de los nidos que cada animal se construye en la copa de los árboles, donde se refugian para dormir. De las raíces de estos pelos se ha extraído ADN y determinado sus huellas características. Al comparar las huellas de los microsatélites de ADN de hembras y machos adultos con los de 13 crías, Gagneux, Woodruff y Boesch han observado que siete crías no habían sido engendradas por machos pertenecientes al grupo. Aunque los investigadores nunca vieron esos encuentros furtivos entre las hembras y machos de bosques vecinos.



**LAS HEMBRAS DE CHIMPANCE, como ésta que aparece aquí con su cría, hacen frecuentes escapadas furtivas para aparearse con machos pertenecientes a otros grupos, según se deduce de estudios con huellas de ADN de microsatélites.**

Estas aventuras nocturnas podrían explicar la diversidad genética mostrada incluso por grupos reducidos de chimpancés. La diversidad refuerza la resistencia frente a enfermedades y favorece, sin duda, la pervivencia.

Es probable que la conservación de la variedad sea esencial para la supervivencia de las poblaciones de chimpancés en libertad. Por desgracia, a medida que estas poblaciones se fragmentan y las distancias que las separan son mayores, la capacidad de las hembras para encontrarse con machos de otros grupos e introducir nuevos genes en sus grupos queda drásticamente recortada.

*—E.R.M. y C.W.*

## Detección del cáncer

El ADN microsatélite puede convertirse muy pronto en un poderoso condicionante de nuestra vida. Con su ayuda podrá afinarse en la detección precoz del cáncer. Las pruebas capaces de descubrir mutaciones en genes cuya alteración predispone al cáncer, pensemos en el *p53* y el *ras*, se emplean ya para detectar hasta una célula maligna en medio de 10.000 normales. Pero las mutaciones en estos genes ni se dan en todos los tipos de cáncer, ni en todos los cánceres de un mismo tipo.

Los microsatélites ofrecen una nueva posibilidad para la detección precoz del cáncer. En efecto, la tasa de expansión o de contracción de microsatélites en las células se dispara en ciertas formas tumorales. Estos cambios repentinos ponen en jaque muchos microsatélites diferentes, fenómeno que puede detectarse con bastante facilidad. Con dicho enfoque se descubre ahora una célula cancerosa entre 500 normales.

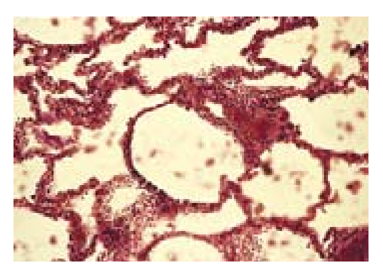
Fue Manuel Perucho quien descubrió en 1993 los cambios operados en los microsatélites de células cancerosas en el marco de una investigación de un tipo de cáncer de colon, hereditario, que no se acompaña de la formación de pólipos. Perucho observó que muchos de los microsatélites de las células cancerosas eran o más largos o más cortos que los de las células normales del mismo paciente. No tardó en demostrarse que uno de los defectos causantes de estas alteraciones se encontraba en un gen determinante de una enzima cuya función era corregir la longitud de los microsatélites que se extienden o menguan durante la replicación del ADN. La pérdida de este gen funcional habría de conducir a un aumento de la probabilidad de que los errores quedaran sin corregir.

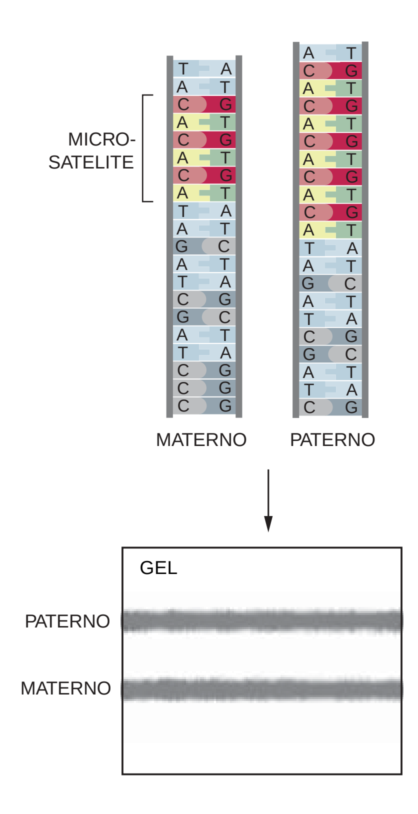
La prueba se completó cuando Richard C. Boland y otros insertaron un cromosoma humano portador de un gen normal reparador de ADN en células cancerosas cultivadas. Observaron que el gen insertado corregía la tendencia a mutar que presentaban los microsatélites de células cancerosas.

Mas, por muy sorprendentes que parezcan estos hallazgos, la inestabilidad de los microsatélites puede constituir un síntoma, no la causa del cáncer. Aunque los ratones en que se ha eliminado el gen que codifica una de las principales proteínas reparadoras de emparejamientos incorrectos de ADN viven poco y adquieren diversos tipos de cáncer, ninguna de las células malignas muestra niveles elevados de mutaciones en los microsatélites. Según parece, este tipo de alteraciones forma parte de los cambios genéticos que se suceden en cascada a lo largo del genoma celular, una vez incoado el proceso de carcinogénesis; ello significa que podría tratarse de productos secundarios del proceso carcinogenético y no de agentes que intervinieran en su desarrollo.

Ahora bien, esas asociaciones se dan con suficiente frecuencia. Y los clínicos empiezan a servirse de la inestabilidad de los microsatélites como un nuevo y poderosos instrumento para la detección precoz del cáncer de colon y de vejiga. Las pruebas clínicas se han visto coronadas por el éxito, lo que explica la voluntad de llevar su aplicación a otros tipos de cáncer. De momento, ninguna de esas pruebas ha salido todavía de los muros de los laboratorios. A medida que los clínicos adquieran experiencia con estos patrones, podrá diagnosticarse el cáncer antes y contar con datos más fiables del tipo de cáncer en razón del patrón de microsatélite.

*—E.R.M. y C.W.*



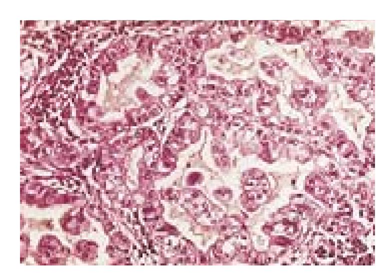




**PRUEBA para la detección del cáncer, que descubre los cambios de longitud de microsatélites; por ejemplo, los formados por repeticiones de la secuencia *CA*. Las células pulmonares normales (*izquierda*) contienen regiones repetitivas de dos longitudes diferentes —una heredada de la madre y otra del padre**

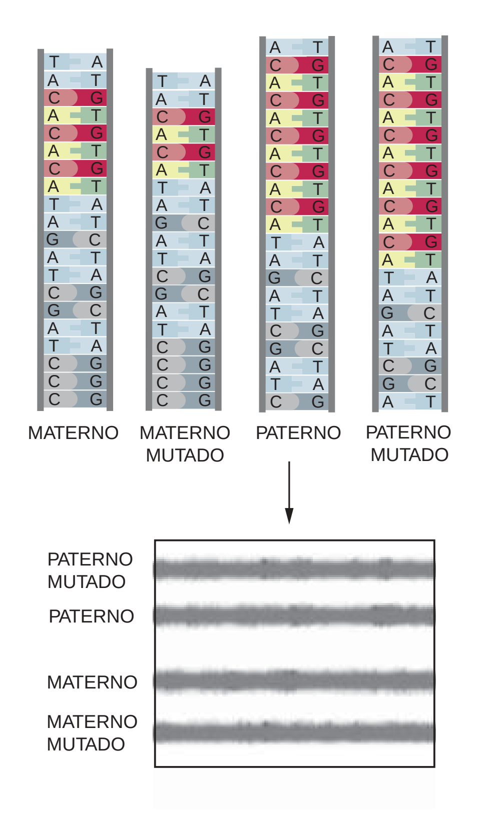
del huésped y otros líquidos tisulares. Las células bacterianas pueden alternar entre los dos estados, según se encuentren sin obstáculos para medrar en el tracto respiratorio o se estén difundiendo por la sangre hacia otros lugares, donde pueden sufrir el ataque del sistema inmunitario.

La mayoría de las bacterias de *H. influenzae* aisladas de humanos son variantes ChoP+, sensibles al ataque inmunitario. Las variantes ChoP– surgen de los emparejamientos incorrectos producidos en el deslizamiento; no persisten habitualmente en el tracto respiratorio, pues se adhieren a las células huésped sin la fuerza de las cepas ChoP+. Ahora bien, si el huésped contrae una infección vírica y se produce una inflamación de la mucosa nasal, se refuerza la exposición de las bacterias a las proteínas defensivas del sistema inmunitario del huésped. En ese caso, las variantes ChoP– tendrían ventaja para repeler el ataque. Una vez que la infección vírica remite, las mutantes ChoP+ generadas por ulteriores emparejamientos defectuosos del ADN microsatélite tornarán a predominar.

Este tipo de genes que se activan o inactivan sin problemas se cobijan bajo la denominación común de genes de contingencia; capacitan, por lo menos a unas cuantas bacterias de una determinada población, para adaptarse a los avatares o riesgos de un nuevo escenario. Entre los



CANCER DE PULMON




**del individuo— que pueden separarse en un gel, de acuerdo con el tamaño. Los tumores tempranos, así el adenoma de pulmón (*derecha*), constan de células normales y células cancerosas, que pueden detectarse porque sus microsatélites se han acortado o alargado y dan más de dos bandas en un gel.**

rasgos codificados por los genes de contingencia no podemos omitir los que gobiernan el reconocimiento inmunitario, la motilidad general, el movimiento hacia las señales químicas (quimiotaxis), la unión a las células huésped y su ulterior invasión, la adquisición de nutrientes y la sensibilidad a los antibióticos. Los genes de contingencia, que representan una fracción muy pequeña del ADN de una bacteria, le confieren una gran flexibilidad en su funcionamiento. Si sólo 10 de los 2000 genes de una bacteria típica son genes de contingencia, por ejemplo, la bacteria podría dar lugar a $2^{10}$, es decir 1024, combinaciones diferentes de genes "activos" o "inactivos". Esta diversidad asegura que al menos una bacteria de la población pueda sobrevivir al ataque inmunitario del huésped o a la acción de otras defensas, para replicarse luego y fundar una nueva colonia.

La inducción de la enfermedad —que se torna en su contra al destruir el huésped que le permite vivir— podría ser parte del precio que las bacterias pagan por disfrutar de esa capacidad de generar tantas variantes. Una de las variantes puede aventurarse allende su nicho ecológico habitual en el tracto respiratorio o el intestino y desencadenar una infección potencialmente letal en otra región del organismo. Habida cuenta de que tal sucede sólo en contadas ocasiones, las ventajas que los genes de contingencia aportan a la pervivencia de una especie bacteriana superan los inconvenientes de destruir algunos huéspedes.

Los microsatélites de estas bacterias son auténticas adaptaciones evolutivas. Es poco probable que secuencias repetitivas tan singulares surgieran al azar; aparecerían, y se han conservado, porque capacitan a las poblaciones bacterianas para una pronta adaptación a los cambios del entorno.

Pese a su interés, diríase que los genes de contingencia se circunscriben a las bacterias. En los eucariotas, nosotros mismos, cuyas células contienen un núcleo, los microsatélites cumplen otras misiones. Ninguno de los microsatélites eucariotas identificados hasta la fecha ha logrado, que se sepa, liar el ADN para que determine proteínas no funcionales. Se sitúan fuera de los genes, salvo un 10 por ciento que cae en su interior. De este 10 por ciento, casi todos pertenecen a los tripletes repetidos, que tienden a extenderse o a contraerse en tríos de bases. Igual que al añadir el artículo "las" o al suprimir el artículo "los" de una frase no se pierde por lo común el sentido, de modo similar estas repeticiones pueden iterarse o perderse sin perturbar el mensaje del gen. Por tener la misma longitud que un codón, podrían dar lugar a la inserción o deleción de unos cuantos aminoácidos repetidos sin que se alterase la secuencia de todos los demás de la ristra.

¿Qué funciones desempeñan, pues, los microsatélites en los organismos superiores? Se sospecha que algunos al menos tienen su cometido. Los eucariotas poseen más microsatélites que las bacterias y muchos de ellos suelen estar alojados cerca o dentro mismo de genes implicados en vías reguladoras de procesos fundamentales. Sin embargo, hasta el momento sólo disponemos de pistas dispersas acerca de su posible tarea.

Los contados efectos cuya pista se ha podido seguir y atribuir a microsatélites eucariotas son, por lo general, nocivos. Así la enfermedad de Huntington, una afección neurodegenerativa grave caracterizada por la aparición tardía de demencia y pérdida progresiva del control motor. Esta patología se dispara con la intervención de un gen defectuoso que codifica una proteína de gran tamaño, la huntingtina, de función desconocida. El gen normal contiene un microsatélite largo, constituido por repeticiones de tripletes, que añade una ristra de glutaminas cerca del extremo inicial de la proteína.

El número de glutaminas del extremo inicial de la huntingtina varía de 10 a 30. Pero las personas con Huntington portan un microsatélite que codifica una ristra de glutaminas excesivamente larga, de hasta 36 unidades o más. Basta la herencia de una copia del gen defectivo —la paterna o la materna— para dictaminar la enfermedad. Se ignora en virtud de qué mecanismo esa ristra de glutaminas repetidas insta la patogénesis.

Hay una docena larga de enfermedades asociadas a tripletes repetitivos; la mayoría son neuropatías raras. Casi la mitad de los microsatélites desencadenantes se alojan en el interior de un gen; en su mayoría, cifran glutaminas. El resto se encuentra cerca de genes y pueden condicionar la función de éstos.

Neuropatía inhabitual es la atrofia muscular bulboespinal, provocada por la extensión de un microsatélite en el interior de un gen del cromosoma X. El gen en cuestión determina un receptor de un andrógeno, hormona sexual masculina. Los individuos con 40 o más repeticiones de estos tripletes en una parte de uno de los genes de receptores de andrógenos acabarán desarrollando la enfermedad. Un grupo dirigido por E. L. Yong ha ido más lejos, demostrando que, cuando estas repeticiones son

ligeramente más largas que lo normal, se presentan complicaciones clínicas. Los varones con entre 28 y 40 repeticiones en la parte del gen del receptor codificadora de glutaminas son con frecuencia infértiles.

Pero una cifra baja de repeticiones en el receptor de andrógenos acarrea también consecuencias adversas. Se ha revelado que los varones con 23 o menos repeticiones son más proclives a padecer cáncer de próstata, aunque se trata de casos bastante raros.

¿Por qué tenemos todas estas bombas de relojería alojadas en nuestro genoma? Es llamativo que las enfermedades relacionadas con tripletes repetitivos afecten al dominio nervioso y ninguno de dichos tripletes se dé en el chimpancé y otros primates. Si este tipo de enfermedades resulta privativo de la especie humana, podría representar el tributo genético que hemos tenido que pagar por la evolución vertiginosa de nuestro cerebro. Es posible que microsatélites largos, en medio de genes determinados o en su vecindad, condicionen la función del cerebro; por cuya razón persistieron quizá durante la evolución, aun cuando se extiendan alguna vez en exceso y sean causa de enfermedad.

Hilvanando un razonamiento teórico, Wills, uno de los autores, propuso en 1989 que algunos genes habrían adquirido capacidad para evolucionar. Según esa hipótesis, en un ambiente que fluctúe de un modo predecible —hacia mayor calor o hacia una situación de frío mayor— la posesión de un mecanismo genético de ágil adaptación tendría ventajas. Los genes de contingencia de las bacterias constituyen ejemplos excelentes de genes con capacidad evolutiva. Merced a una frecuente mutación hacia delante y hacia atrás, las bacterias se adaptan presto a cambios ambientales predecibles y revierten a las condiciones previas cuando reaparece la situación de entonces.

Tal vez los microsatélites eucariotas ejercen una forma más sutil de regulación que la que proporcionan a las bacterias sus genes de contingencia. En el hombre los microsatélites alojados en medio de genes afectan a la velocidad de producción de ciertas proteínas, desde la que acompaña al pigmento biliar bilirrubina a determinados neurotransmisores, sustancias químicas que transmiten mensajes entre neuronas. Para David G. King, estos microsatélites constituirían una suerte de "botones de ajuste fino", que habrían evolucionado actuando como reostatos de la función del gen, reforzando la síntesis de una proteína en unos casos o frenándola en otros.

El grupo de Walter Schaffner ha demostrado que la adición de microsatélites que añaden hileras de glutaminas o prolinas (otro aminoácido) en el comienzo de un determinado gen puede incrementar su capacidad para producir proteína. Tal vez, por ser menos perturbadora que la conmutación de los genes de contingencia, esta forma de regulación génica emergió durante la evolución de organismos pluricelulares complejos.

Han comenzado los primeros sondeos sobre las funciones desempeñadas por los microsatélites en nuestra especie. Estas secuencias repetitivas, con su capacidad de cambiar rápidamente entre un número limitado de estados, nos ofrece una pista nueva para conocer nuestro poder de adaptación al cambio ambiental, a imagen de lo realizado por los genes de contingencia en bacterias.

BIBLIOGRAFIA COMPLEMENTARIA


ADAPTIVE EVOLUTION OF HIGHLY MUTABLE LOCI IN PATHOGENIC BACTERIA. E. R. Moxon y cols. en *Current Biology*. vol. 4, n.º 1, páginas 24-33; 1 de enero de 1994.

CANCER OF THE MICROSATELLITE MUTATOR PHENOTYPE. M. Perucho en *Biological Chem.*, vol. 377, págs. 675-684, 1996.

EVOLUTIONARY TUNING KNOBS. D. G. King, M. Soller e Y. Kashi en *Endeavour*, vol. 21, n.º 1, págs. 36-40; 1997.

FURTIVE MATING IN FEMALE CHIMPANZEES. P. Gagneux, D. S. Woodruff y C. Boesch en *Nature*, vol. 387, págs. 358-359; 22 de mayo de 1997.

SOMATIC FRAMESHIFT MUTATIONS IN THE BAX GENE IN COLON CANCERS OF THE MICROSATELLITE MUTATOR PHENOTYPE. N. Rampino, H. Yamamoto, Y. Ionov, Y. Li, H. Saway, J. Reed y M. Perucho en *Science*, volumen 275, páginas 967-969; 1997.

CÁNCER DEL FENOTIPO MUTADOR DE MICROSATÉLITES, Manuel Perucho en *Investigación y Ciencia*, n.º 261, págs. 46-55; junio de 1998.

# El problema del año 2000

*La solución del problema del año 2000 no parece difícil: bastará representar con cuatro cifras los años expresados con dos. Pero esa tarea, tediosa y compleja, requiere más tiempo del disponible*

Peter de Jager

Parte de la explicación de que el año 2000 provoque un problema informático pertenece al dominio de la psicología. Se halla escondida en un lugar inesperado, en el carrolliano *Alicia en el País de las Maravillas*, pieza clásica de la literatura infantil. Allí, el Sombrerero Loco pregunta: "¿Dice tu reloj en qué año estamos?" Y Alicia responde: "¡Claro que no! Pero eso es porque sigue siendo el mismo año mucho tiempo."

Muchas han sido las causas de que los programadores, el autor entre ellos, optasen por representar los años con sólo dos dígitos; por ejemplo, 55 para expresar 1955, o 23/10/76 para el 23 de octubre de 1976. Hace decenios, los soportes materiales de la informática eran muy costosos y los recursos, escasos. La memoria de los ordenadores era muy cara. Las tarjetas perforadas típicas tenían una anchura de sólo 80 columnas. En favor de la notación abreviada se adujeron también razones de eficiencia, por ser menor el número de pulsaciones a efectuar en los teclados. La inexistencia de notaciones normalizadas consintió que estas ideas se llevaran a la práctica. Y por otra parte, muchos de nosotros creíamos sincera (aunque erróneamente) que los programas que estábamos redactando habrían quedado fuera de servicio mucho antes de la conclusión del milenio. La pura inercia, sumada al argumento de Alicia en el té del Sombrerero (¿para qué más de dos cifras, si el siglo dura tanto?), hizo que se continuase actuando igual, mucho después de que las consideraciones de memoria y coste hubieran dejado de ser razones válidas para ello.

¿Cuál ha sido el resultado? Los ordenadores están plagados de representaciones de fechas de una ambigüedad pavorosa. Con pocas palabras, ¿cómo sabrán las máquinas digitales si 00 significa 1900 o 2000? Tal confusión ha suscitado ya no pocos problemas. Allá por 1993, la empresa Boeing observó errores en una aplicación que gestiona con antelación de siete años los encargos efectuados. Uno de los sistemas de Amway Corporation, una firma de jabones y artículos de higiene y cosmética extendida por todo el mundo, rechazó ciertos productos químicos al considerar, erróneamente, que tenían casi un siglo de antigüedad. Y algunos sistemas de cajas registradoras computarizadas se han quedado "en blanco" cuando los clientes intentaron utilizar tarjetas de crédito que caducaban en el año 00. De hecho, según un estudio llevado a cabo hace un año en la industria, más del 40 por ciento de las organizaciones estudiadas habían padecido ya algún tropiezo con "el problema del año 2000", o en la jerga, el A2K.

## La esencia del problema A2K

¿Dónde se halla el origen de tales dificultades? Tendríamos un ejemplo obvio en una cuenta de ahorro abierta en 1999 y saldada en el año 2000. Si el programa de contabilidad del banco calcula los intereses generados restando 99 de 00, el ordenador considerará, erróneamente, que la duración del depósito ha sido de –99 años.

Veamos un yerro de tipo A2K algo menos palmario. Una compañía de seguros se dedica rutinariamente a buscar pólizas inactivas durante más de cinco años, para eliminarlas de sus bases de datos. Al objeto de localizar las pólizas inactivas, la firma hace funcionar una aplicación que actúa sobre datos almacenados bajo el concepto ULTIMO-ACCESO, en el que están consignadas las fechas más recientes en las que los asegurados efectuaron alguna transacción.

Cuando el programa examina un valor de ULTIMO-ACCESO, le suma al año, expresado con dos dígitos, 5 unidades. Si la suma resultante es menor que el año en curso, procede a eliminar el registro. Así pues, si ULTIMO-ACCESO hubiera sido 93, se tendría $93 + 5 < 99$, y la póliza sería, correctamente, declarada inactiva. Sin embargo, de haber sido 96 su valor, el procesamiento se vuelve incierto. Si sólo se permiten representaciones con dos dígitos, siendo truncada la información restante, 96



**FALLOS YA OCURRIDOS**

***La compañía de seguros Unum Life eliminó 700 registros de una base de datos que controlaba el estado de las licencias de agentes de bolsa y cambio, porque un ordenador confundió 00 con 1900.***

***Mary Bandar, una centenaria, recibió una invitación para matricularse en un jardín de infancia de Winona (Minnesota).***

***En 1993, Boeing experimentó errores en un sistema que utilizaba siete años de antelación para los encargos.***

***Un sistema de mezclado de Amway Corporation, controlado por PC, rechazó un lote de productos químicos porque consideró, erróneamente, que su fecha de caducidad era 1900.***

PETER DE JAGER se dedica a la asesoría técnica. A su ensayo "Doomsday 2000", publicado hace más de cinco años en el semanario *Computerworld*, se le atribuye haber alertado a la industria y a los gobiernos sobre el problema informático del AK2. De Jager es coautor de *Managing 00: Surviving the Year 2000 Computering Crisis* y de *Countdown Y2K: Business Survival Planning for the Year 2000*. Hace algunos años, se le acusó de alimentar recelos tecnófobos; hoy es tenido por moderado ante tanto catastrofista del A2K.

+ 5 se convierte en 01, que es menor que 99. El resultado sería la incorrecta eliminación de la póliza. (Como este ejemplo es revelador de una programación chapucera, nos sentimos tentados a no tenerlo en cuenta. Sin embargo, fallos de este tipo han ocurrido ya.)

En ciertos casos, un error A2K podría tener, literalmente, consecuencias mortales. Un médico, al utilizar un programa clínico que especifica la posología de ciertos fármacos, podría anotar "16-03-00" al dar la fecha de nacimiento de un niño. El ordenador, sin embargo, podría dar por supuesto que el paciente cuenta 100 años y recomendar una dosis idónea para un adulto provecto, pero letal para un bebé.

Tal es, pues, la esencia del problema del A2K. Al registrar las fechas, se han consignado de forma ambigua los años. Las confusiones consiguientes comienzan a crear un infierno. Muchas organizaciones se han limitado a sustituir los programas conflictivos por otros nuevos en los que ya está resuelto el problema. En el caso de sistemas "empotrados", esa táctica ha resultado muchas veces la única opción viable. Pero la sustitución por otros nuevos de la miríada de programas hoy en uso sería excesivamente onerosa y exigiría demasiado tiempo. Por eso, las compañías se aprestan a remendar lo que tienen.

**Pero, ¿qué es una fecha?**

En cuanto se tomó conciencia del alcance del problema A2K, se propuso una estrategia de reparación válida e incluso óptima: si lo hecho ha sido eliminar dos dígitos, repongámoslos. Dejando aparte por el momento las complejidades logísticas que entraña la modificación de los teraoctetos de datos de cualquier gran sociedad anónima multinacional, esta solución plantea una cuestión crucial, por peregrina que parezca: ¿qué es una fecha? Después de todo, para añadir dos dígitos a cada fecha, habrá que encontrarla antes.



Pero los ordenadores no saben absolutamente nada de lo que en realidad representa

la información que manipulan. Somos nosotros quienes concedemos significado a los datos; con mayor precisión, el significado de los datos está en nosotros. Y al crear los nombres con los que identificar fechas, los programadores no se atuvieron a normas fijas. Las fechas pueden tener nombres tan poco expresivos como SNARK o WUMPUS, aparte de elecciones más obvias, como FECHA_NAC y FCH_ALTA_EMP.

Tampoco resulta fiable la información numérica propiamente dicha. Uno de los métodos aplicados ha consistido en ir filtrando los datos, seleccionar los que contengan columnas con valores de 1 a 31, de 1 a 12, y de 0 a 99; para presumir, por fin, que el último elemento de información es denotativo de años. Esta presunción, aunque explicable, no es siempre correcta. El dato del intervalo 0-99 podría representar porcentajes de cierta magnitud. Y no es raro que los dos dígitos correspondientes a los años se encuentren profundamente sepultados en el seno de otros datos, como ocurre en los largos números de serie de muchos productos.

Se han confeccionado instrumentos informáticos para la localización automática de fechas. Algunos de los productos más avanzados son capaces de razonamientos deductivos, y alcanzan índices de éxito asombrosos. Mas no existe todavía ningún instrumento que haya logrado un funcionamiento impecable, sin ningún error.

El hallazgo de las fechas es el primer paso. Ahora bien, en la ampliación a cuatro cifras de la notación de los años los programadores pueden verse obligados a rediseñar el formato de ciertos informes, formularios y pantallas de ordenador para dar cabida a los dígitos añadidos. Otra complicación de mayor alcance es que podría también requerirse modificar las aplicaciones informáticas que hacen uso de los datos modificados.

Imaginemos, por ejemplo, el caso de la base de datos del personal de una empresa, donde las columnas 1 a 30 de los registros están reservadas para nombres de los empleados; para sus fechas de nacimiento, las columnas de la 31 a la 36; para los salarios, las columnas 37 a 42, y así sucesivamente. Si ampliamos el espacio de las fechas de nacimiento para dar alojamiento a los cuatro dígitos de los años, la información sobre salarios con todos los datos subsiguientes tendrá que ser desplazada dos columnas hacia la derecha. En consecuencia, los programas que recurran a esa información habrán de sufrir retoques, para que tomen los datos de las posiciones adecuadas.

De ordinario, tales modificaciones se efectúan sobre el "código fuente", que los programadores redactan en lenguajes COBOL y C. El programa se convierte después a módulo objeto, una forma susceptible de ser ejecutado por un ordenador. Dicha traducción se efectúa mediante programas "compiladores". Pero pueden surgir problemas, pues los compiladores se depuran y actualizan sin cesar. En algunos casos, técnicas de programación que compiladores anteriores aceptaban como legítimas podrían estar prohibidas ahora, de forma muy parecida a como un documento producido por la versión 3 de un programa de proceso de textos podría ser rechazado por la versión 5. Por ello, el antiguo código fuente, recién reparado para el A2K, puede no quedar correctamente compilado y convertido en nuevo módulo objeto, a menos que se efectúen modificaciones adicionales.

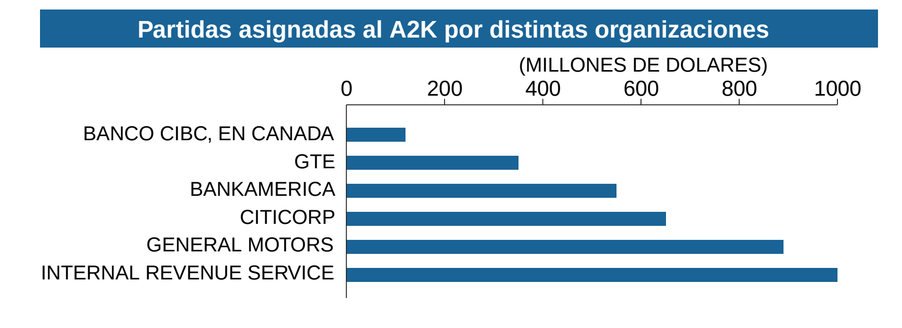
Para agravar las cosas, muchas compañías han perdido parte del código fuente de sus programas. Aunque la cantidad de código perdido no suele llegar a un 3 o un 4 por ciento, esa pequeña cantidad resulta fastidiosa en extremo, porque los programadores no pueden modificar fácilmente un módulo objeto. Tienen que volver a crear el código fuente, sea a partir de cero (tarea difícil, porque lo más probable es que se carezca de la documentación correspondiente a la parte desaparecida) o partir del propio módulo objeto (proceso muy penoso, que se ha comparado con el de reconstruir un cerdo a partir de una salchicha).



**LOS PROGRAMADORES han estado muy ocupados solventando problemas A2K. En esta "factoría" de programas de Charlotte (Carolina del Norte) más de 250 personas de la plantilla de la empresa Alydaar Software reparan cada mes millones de líneas de código. En su mayoría las empresas contratantes de ese servicio figuran en Fortune 500.**

**LITIGIOS DEL A2K**

***Después de que las cajas registradoras automáticas de Produce Palace, en Warren (Michigan), se colgasen una y otra vez cuando las tarjetas de crédito de los clientes expiraban en 00, el comerciante demandó ante los tribunales a TEC America, empresa fabricante del sistema.***




**OTRAS FECHAS DE RELANZAMIENTO**

***22 de agosto de 1999: El sistema de localización global (GPS) reanudará su ciclo desde la semana 0, al cabo de 1023 semanas.***

***19 de enero de 2038: Los sistemas Unix reanudarán su ciclo, porque habrán transcurrido $2^{31}$ segundos desde la fecha de haber sido puestos en servicio, el 1 de enero de 1970.***

***6 de febrero de 2040: Los ordenadores Macintosh más antiguos pueden retornar al 1 de enero de 1904, perdiendo $2^{32}$ segundos.***

Una vez reparado y recompilado el código fuente, es preciso ensayarlo. Dado que en las revisiones de los programas casi siempre se deslizan nuevos errores, ocurre que la verificación de los programas modificados constituye en la actualidad la mayor parte del proyecto Año 2000.

A principios del decenio en curso muchos expertos sostuvieron que el método más adecuado para afrontar el A2K sería la ampliación de los campos de fechas. Pero la inesperada necesidad de tener que recopilar todos los programas que aludan a alguna fecha de algún archivo, aun cuando la aplicación no realice cálculos con fechas, ha hecho que tal método les resulte demasiado caro y lento a las empresas. (Evidentemente, la ampliación a cuatro dígitos provocará, en la transición del año 9999 al 10.000 d.C., un problema similar, el problema A10K. Pero eso es tema de otro artículo.)

### Abrir ventanas

Una solución de otro tipo consiste en enseñar a los ordenadores que 00 significa 2000. Los programadores han generalizado esta idea, estrategia que en su jerga denominan "enventanar". Han tomado todos los años, desde 00 hasta 99, y los han repartido en dos grupos atendiendo a un valor de divisoria, o valor de umbral, cuidadosamente elegido (por ejemplo, 45). Los años que expresados con dos dígitos sean mayores o iguales que el umbral se consideran pertenecientes al actual siglo XX (68 se convierte en 1968). Todos los demás, pertenecientes al siglo XXI (13 pasa a ser 2013).

Valiéndose de esta idea, el programador puede examinar el código fuente y hallar todas las referencias a fechas, para modificar en consecuencia los cálculos que hagan uso de esa información. Puesto que no se modifican los años consignados con dos dígitos, la técnica de "ventanas" requiere menos trabajo que la ampliación de fechas, y es en la actualidad la más utilizada para resolver el A2K. (Lo único que se modifica son los cálculos en los que intervienen las fechas, para situar los años en los siglos correctos.)

Pero el método no puede ocultar sus puntos flacos. Falla, según es obvio, para datos que abarquen intervalos de más de 100 años, como ocurre con las fechas de nacimiento y los alquileres a muy largo plazo. Plantea, asimismo, ciertos problemas interesantes cuando es preciso trasladar información entre sistemas que han utilizado valores umbrales distintos. Por ejemplo, para un programa encargado de las facturas de una empresa fundada en 1928, podría ser adecuado situar la divisoria en 25. Pero otra aplicación que efectúe previsiones a largo plazo de ventas para esa misma empresa podría situar el umbral en 70. El problema, evidentemente, se plantea cuando es preciso trasladar información de uno a otro programa: el año 2031 se convertirá en 1931, o viceversa, provocando un caos en los cálculos.

Para complicar las cosas, los programadores han utilizado "ventanas corredizas", donde el valor umbral, la divisoria, va cambiando con el tiempo. Esta estrategia queda justificada en ciertos tipos de programas; por ejemplo, los de gestión de hipotecas a 30 años. En aplicaciones de este tipo, la divisoria podría estar en 40 años contados a partir del actual. Ni que decir tiene que debe someterse a un control estricto las divisorias móviles, para evitar posibles conflictos con otros programas. Los programadores, además, han de percatarse de que la representación del año 1968 en un programa con divisoria en 70 exige nuevos retoques en los programas.

### Traslación cronológica

Otro de los métodos utilizados para afrontar el A2K se vale de aritmética elemental. Fijémonos en el cálculo 00 – 99 = –99. Si esta operación se proponía representar la diferencia 2000 – 1999, es obvio que la solución –99 es incorrecta. Observemos, empero, que 00 – 99 es equivalente a (00 + 5) – (99 + 5). Si esta expresión se evalúa en aritmética de dos dígitos (o módulo

**¿Está preparado el gobierno federal estadounidense?**

| | Número total de sistemas críticos | Sistemas reparados en agosto de 1998 | Sistemas reparados en marzo de 1999 (estimación) | Año en que todos los sistemas estarán reparados (estimación) |
|---|---|---|---|---|
| SSA<br>Administración de la Seguridad Social | 308 | 93% | 99% | 1999 |
| FEMA<br>Agencia Federal de Gestión de Emergencias | 49 | 69% | 82% | 1999 |
| Hacienda<br>Departamento del Tesoro | 323 | 45% | 61% | 2000 |
| DOD<br>Departamento de Defensa | 2965 | 42% | 54% | 2001 |
| Justicia<br>Departamento de Justicia | 207 | 31% | 31% | 2030+ |
| NASA<br>Administración Nacional de Aeronáutica y del Espacio | 158 | 63% | 74% | 2000 |

## Problemas empotrados en todas partes

Ocultos en el interior de un equipo utilizado en tratamientos médicos con radiaciones hállanse microcircuitos de cómputo, encargados de garantizar que a cada paciente se le aplica la dosis correcta. Pero, dado que los cálculos se efectúan con expresiones de dos dígitos para los años, cuando se pase del siglo actual al siguiente el aparato no podrá calcular correctamente la intensidad de la radiación de una muestra en desintegración. Pero el fabricante ha detectado a tiempo el problema y se ha puesto en contacto con los hospitales para retirar el producto.

Sin duda, los comodines más peligrosos del A2K son microcircuitos preprogramados e instalados en equipos electrónicos muy diversos: sistemas de maquinaria industrial, dispositivos de control o supervisión, semáforos, sistemas de alarma, instrumentos de navegación y un sinfín de bienes de consumo, como automóviles, relojes, magnetoscopios y hornos de microondas. Estos "sistemas empotrados" se encuentran por doquier, de las fábricas a las centrales nucleares, pasando por los sistemas de suministro de aguas y de saneamiento, hospitales, edificios de oficinas y domicilios particulares. El Grupo Gartner, una asesoría de Stamford, Connecticut, sitúa el total mundial entre 32 y 40 millardos.

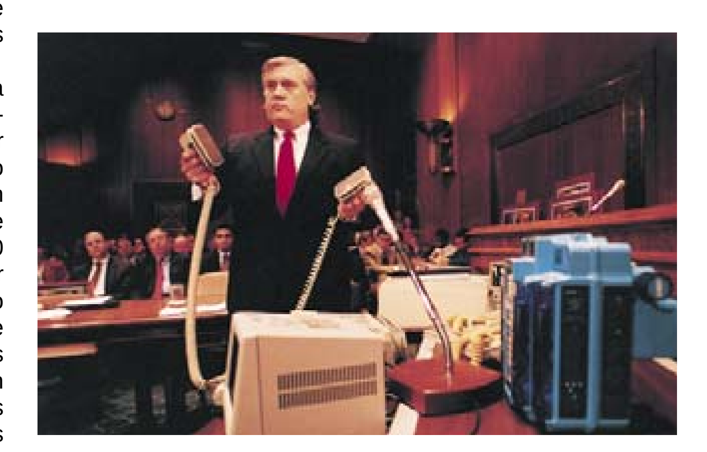
La mayoría de estos sistemas no dependen del conocimiento de fechas; no se verán afectados por el A2K. Entre los sensibles a las fechas, sólo una fracción muy diminuta presentará anomalías de procesamiento. Pero basta esa mínima fracción para preocuparse. Nadie sabe de cierto su número exacto, pero Gartner cifra en millones los sistemas empotrados vulnerables.

La existencia de tan enorme incertidumbre se debe a que resulta dificilísimo, incluso para ingenieros experimentados, determinar con exactitud de qué modo podría fallar un sistema empotrado. A menudo, equipos que en principio no dependen de cómputos con fechas, sí dependen en realidad. Imaginemos un dispositivo que por razones de seguridad se desconecta por sí solo el 1 de enero de 2000 considerando, erróneamente, que fue inspeccionado por última vez hace casi un siglo. Un problema más insidioso aún podría presentarse con un sistema de supervisión que emitiera advertencias de alarma. Los mensajes cruciales podrían ignorarse porque, al llevar fecha 00, desaparecen de la pantalla del operador desalojados por otros mensajes que, con fecha 99, parecen más recientes. La verdad es que el A2K puede suscitar problemas en los lugares más inverosímiles. Los bomberos de Baton Rouge (Louisiana) tuvieron que probar sus camiones cisterna para averiguar si los mecanismos de control de las bombas de agua y las escalas requerían reparaciones A2K. Por fortuna no es así.

Los ordenadores personales figuran entre los productos en que la corrección de problemas empotrados es más sencilla. La mayoría de los PC adquiridos antes de 1996 no saben del cambio de siglo: el 1 de enero de 2000 retornarán automáticamente a su fecha de reposición, en 1980. Pero existen ya "parches" que corregirán automáticamente este fallo.

Otros arreglos son menos sencillos. Muchas de las compañías fabricantes de sistemas empotrados ya no existen; si sobreviven, es probable que los diseñadores de los productos defectuosos no trabajen ahora en ellas y que la documentación del producto se haya perdido o esté incompleta. A mayores, algunos de los sistemas deficientes se encuentran instalados en lugares remotos, como las plataformas petrolíferas en el mar.

Una solución común consiste en la sustitución uno por uno de los microcircuitos defectuosos, entre ellos los relojes en tiempo real (que controlan el paso del tiempo mediante un oscilador de cuarzo) y los microprocesadores y controladores (que ordenan a un dispositivo la ejecución de determinadas acciones). En ciertos casos, el proceso se reduce a extraer de los circuitos impresos los componentes referidos e implantar otros adaptados. A menudo habrá que desechar el equipo completo y reemplazarlo por otro nuevo, lo que sin duda es costoso, pero a veces inevitable, en especial si el aparato no es capaz de calcular correctamente las dosis de radiación de los pacientes.



**KENNETH W. KIZER, de la Administración Sanitaria de Veteranos, en su testimonio ante el Congreso en julio del año pasado, manifestó que ciertos equipos clínicos, entre ellos desfibriladores cardíacos, no funcionarían correctamente por culpa de problemas derivados del fenómeno A2K en sus microcircuitos.**

100) el resultado será 5 – 4 = 1. De hecho, la adición de 5 a 00 y a 99 ha trasladado ambas fechas a un mismo siglo, con lo que el cómputo de 2000 – 1999 podría haberse efectuado correctamente con dos dígitos.

Pero las fechas son más que meros números. El 1 de enero de 2000 caerá en sábado; el 1 de enero de 2005, no. Así pues, la solución consistente en sumar 5 a todos los años fallará en los programas que deban distinguir los días de la semana. Pero no todo está perdido. Básicamente, los días de la semana se rigen por dos ciclos: un ciclo de período siete para los días; un ciclo de período cuatro, para los años bisiestos. Al multiplicar ambos se obtiene un período conjunto de 28 años. El 1 de enero de 2000 será sábado, como lo fue el 1 de enero de 1972, y como lo será el 1 de enero de 2028. La técnica de "encapsulado" saca provecho de esta regularidad, y suma 28 años a los dos dígitos de años

### EL A2K EN JAPÓN

***Japón ha recibido críticas por su pasividad ante el problema A2K. Pero uno de los aspectos en favor de ese país es que muchas compañías han utilizado el calendario del sistema imperial, en el cual 1999 es Heisei 11.***

## La situación del A2K en los EE.UU.

| | Caso óptimo | Caso esperado | Caso pésimo |
|---|---|---|---|
| Aplicaciones informáticas con problemas A2K | 10 millones | 12 millones | 15 millones |
| Porcentaje de problemas A2K no resueltos a tiempo | 5% | 15% | 25% |
| Fallos en infraestructura debidos al A2K* | | | |
| Plantas de generación de energía | 5% | 15% | 75% |
| Sistemas de transporte | 5% | 12% | 50% |
| Sistemas de telefonía | 5% | 15% | 65% |

*Porcentaje de hogares que serán afectados

antes de efectuar cálculo alguno. Una vez llevados a cabo los cálculos, se restan 28 de ambas fechas.

Aunque el encapsulado pueda servir para eludir muchos problemas de A2K, esta técnica resulta de difícil manejo en cómputos complejos. Tomemos, por ejemplo, el caso de una aplicación que opera en paralelo con uno o varios programas más: durante sus propios cálculos, la aplicación puede tener que enviar información a otros sistemas que manipulen los datos según sus propios procedimientos; los resultados podrían después combinarse y recombinarse en diversos momentos. Los programadores tendrían dificultad para determinar y controlar en qué casos se ha de sumar y en cuáles restar 28.

El encapsulado falla, además, en el caso de fechas inmersas en otra información que utiliza ciertos dígitos con fines de validación. Tomemos un número de serie de un producto, como 7289-47-99-5, en el cual 99 representa la fecha de caducidad; el último dígito, 5, se utiliza como dígito de verificación de la secuencia completa. En este caso, el 5 resulta de efectuar la suma de 7289, 47 y 99, y tomar la última cifra del resultado, 7435. Estos dígitos de comprobación suelen emplearse para verificar los números de tarjetas de crédito, códigos de barras y números de la Seguridad Social. Como es obvio, si se suma alegremente 28 a las representaciones bidígitas de los años se echarían a perder estos cálculos de validación.

Las tres técnicas, a saber, la ampliación de los campos de datos, el enventanado y el encapsulado, han dado cuenta hasta el momento de más del 95 por ciento de las reparaciones por A2K efectuadas a programas existentes. Muchas compañías grandes, que cuentan con millares de programas informáticos, han adoptado métodos mixtos. Existen útiles informáticos para automatizar las tres soluciones, pero ninguno ha conseguido un éxito completo, sin error alguno.

No deja de ser perverso que uno de los factores que dificultan la adaptación al A2K sea el que algunos programas han tenido ya en cuenta el cambio de siglo. En previsión de los problemas A2K, sus programadores han empleado sus propios sistemas de ventanas o encapsulados. Pero puede ocurrir que estas precauciones no se hayan tomado en la totalidad de la programación, dejando al sistema a medio camino hacia una solución. En tales programas, la implantación de remedios adicionales podría provocar un embrollo digital: dobles ventanas con umbrales en distinto año o aplicación a unos mismos datos de una combinación de encapsulado y ventanas.

La averiguación de si en un programa se han efectuado arreglos previos se torna ardua si le falta pulcritud a la codificación o se opera con una documentación deficiente. Me he encontrado, como programador que soy, sentado durante horas frente a un breve fragmento de código (sólo 10 o 15 líneas, tal vez), tratando de averiguar cuál era su finalidad. Y el peligro de los útiles automáticos es que aplican ciegamente una solución estereotipada, sin comprender lo que en realidad pretende hacer el código.

Existen varios obstáculos más. Ha sido costumbre, al programar, conferir a las fechas o años próximos al 2000 un significado especial. En concreto, ha sido frecuente usar la marca 9999, o sencillamente 99, para señalar el final de un archivo o un registro que debía suprimirse o archivarse. Esa costumbre induce a confusión, pues las dos cantidades podrían legítimamente significar 9 de septiembre de 1999, o el año 1999, respectivamente. Por ejemplo, una aplicación de ventas podría pedir a los empleados que introdujeran la clave 99 como si fuera un año si quisieran eliminar el correspondiente encargo de un cliente. El programa ha de reescribirse *ex novo* para posibilitar que la petición se efectúe de otro modo.

Los años bisiestos entrañan su complicación. Dado que la Tierra invierte un poco menos de 365,25 días en su órbita alrededor del Sol (un valor más exacto es 365,242199), los años bisiestos no siguen estrictamente un ciclo cuatrienal. Los años excepcionales son los denotativos de fin de siglo, por ejemplo, 1700 y 1800, que no son bisiestos; a su vez, éstos tienen como excepciones los siglos divisibles por 400. En consecuencia, habrá un 29 de febrero en el año 2000 (a pesar de que no lo hubo en 1900), creando todavía mayor confusión en los ordenadores que no hayan sido debidamente programados con esa información.

### ¿Se avecina un desastre digital?

Ese manojo de factores, entre otros muchos, han suscitado fuertes incertidumbres y acaloradas controversias acerca del A2K. En un extremo del espectro se oyen las simplezas más absurdas: "Existe la posibilidad de que nos quedemos sin

**ESCASEZ DE CIRCULANTE**

***En previsión de que la gente haga acopio de dinero en efectivo, la dirección de la Reserva Federal estadounidense tiene la intención de aumentar en 50 millardos de dólares la masa dineraria circulante.***



energía eléctrica para siempre...” (opinión manifestada hace poco por uno de los ponentes en una conferencia sobre A2K). En el otro dormita la complacencia desinformada: “El A2K es flor de un día. Los problemas que pueda plantear se arreglan en un fin de semana.”

Ambos extremos revelan la misma ingenuidad. La primera desprecia la capacidad de la sociedad para recuperarse tras la adversidad. La idea de que por alguna razón los humanos van a perder la capacidad intelectual para producir electricidad no merece una línea. De hecho, las organizaciones que, como las instituciones financieras, han dedicado los recursos necesarios, han conseguido grandes progresos al respecto del A2K. El verano pasado, Wall Street efectuó una simulación de cómo se desarrollaría la sesión del 3 de enero de 2000 y sólo detectó problemas de poca monta. Se prevén nuevas pruebas para la primavera del año en curso.

Por otra parte, desdeñar el A2K equivale a cerrar los ojos ante la vulnerabilidad técnica de la sociedad moderna, que se sostiene sobre intrincados cimientos de interdependencias mutuamente imbricadas. En particular, los efectos de las fracturas en puntos concretos podrían expandirse rápidamente por todo un sistema, con resultados desastrosos. En la primavera pasada, el *Galaxy IV*, uno de los muchos satélites de comunicaciones, sufrió una avería y millones de buscapersonas sufrieron una muerte súbita. En Auckland, Nueva Zelanda, falló un cable eléctrico, sólo uno. La red de suministro eléctrico experimentó una sobrecarga, y la ciudad malvivió seis semanas con cortes eléctricos.

Se trata de hechos reales, ninguno de los cuales estaba previsto. Pero el caso del A2K es diferente. Conocemos su venida. En nuestros días, profesionales informáticos de todo el mundo están modificando gran parte de los programas existentes. En Canadá, el banco CIBC tiene trabajando en el proyecto a 1000 personas, con un presupuesto de unos 120 millones de dólares. La corporación AT&T ha invertido ya más de 500 millones de dólares, Citicorp va a dedicar alrededor de 650 millones, y el IRS estadounidense (Internal Revenue Service, homólogo de la Agencia Tributaria española) gastará en torno a 1000 millones de dólares.

Aunque se trata de esfuerzos enormes, si alguna lección han enseñado los proyectos informáticos grandes es que son muchos los no concluidos en los plazos previstos, y que los sí terminados rara vez funcionan a la perfección. Negar esa realidad supone olvidar las enseñanzas de estrepitosas debacles del pasado, como el fiasco de los Juegos Olímpicos de Atlanta o del Aeropuerto Internacional de Denver. No es frecuente la instalación a tiempo y sin errores de sistemas informáticos complejos. Lo insoportable de los proyectos A2K es la inamovilidad de la fecha fatídica.

Dicho todo lo anterior —y tenidos en cuenta el volumen de trabajo realizado, las contingencias previstas y las componendas que será necesario aceptar al aproximarse la vuelta de siglo—, estoy convencido de que se producirán quiebras importantes de funcionamiento cuya reparación costará días. Otros problemas, desde pequeñas molestias hasta tropiezos en asuntos serios, irán apareciendo a lo largo del año 2000. Quizá peque de optimista, al dar por supuesto que se ha hecho lo necesario para minimizar el número de puntos de ruptura que pudieran surgir. La consecución de este objetivo en el tiempo disponible va a exigir un esfuerzo hercúleo, sin precedentes en la historia de la informática.

BIBLIOGRAFIA COMPLEMENTARIA


YEAR 2000 PROBLEM: STRATEGIES AND SOLUTIONS FROM THE FORTUNE 100. Leon Kappelman. International Thomson Computer Press, 1997. ISBN: 1-85032-913-3.

THE YEAR 2000 SOFTWARE CRISIS: CHALLENGE OF THE CENTURY. William M. Ulrich y Ian S. Hayes. Prentice Hall, 1997.

COUNTDOWN Y2K: BUSINESS SURVIVAL PLANNING FOR THE YEAR 2000. Peter de Jager y Richard Bergeon. John Wiley & Sons, 1998.

EVALUATING SUCCESS OF A Y2000 PROJECT. Howard Rubin y Brian Robbins. Information Economics Press, 1998.

YEAR 2000 COMPUTER CRISIS: LAW, BUSINESS, TECHNOLOGY. Michael D. Scott y Warren S. Reid. Glasser Legal Works, 1998.

THE YEAR 2000 SOFTWARE PROBLEM: QUANTIFYING THE COSTS AND ASSESING THE CONSEQUENCES. Capers Jones. Addison-Wesley, 1998.

Existe más información sobre el A2K en el nodo Web del autor: www.year2000.com

Shawn Carlson

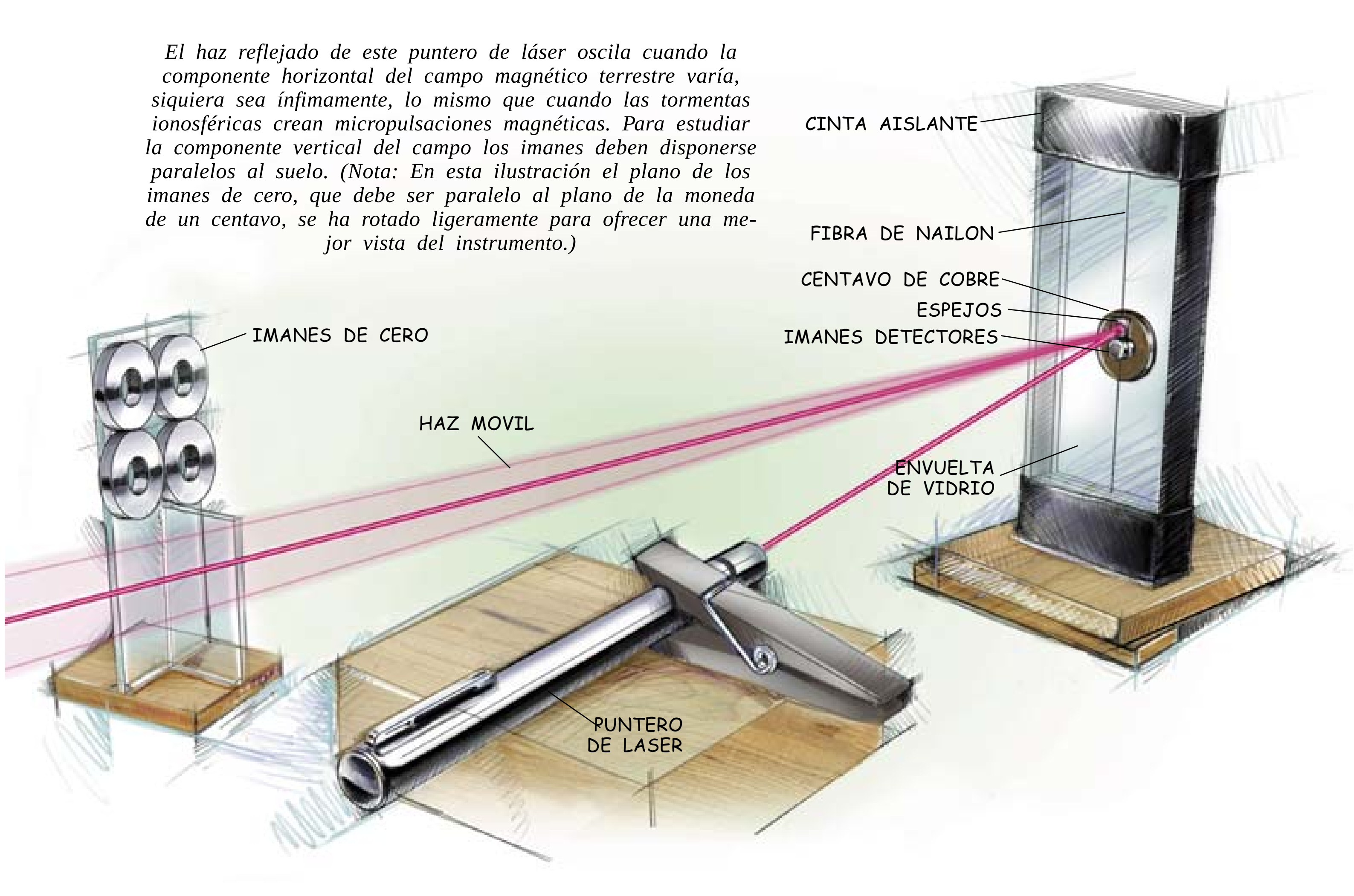
# Tomando el pulso magnético a la Tierra

La gente sólo piensa en el magnetismo terrestre cuando está perdida. Si miramos la brújula, el campo magnético se nos ofrece como un guía estable. La verdad es que el campo magnético es cualquier cosa menos estable. Efímeras ondulaciones llamadas micropulsaciones rizan la ionosfera y generan perturbaciones magnéticas que descienden hasta el nivel del suelo. Aunque sean un fenómeno habitual que a veces dure entre segundos y minutos, esas perturbaciones son de detección difícil y su intensidad apenas llega a la diezmilésima de la del magnetismo terrestre medio.

Durante decenios, el elevado precio de los magnetómetros de buena sensibilidad ha reservado el seguimiento de esas señales a los profesionales. Pero ahora, gracias al genio creativo de Roger Baker, de Austin (Texas), el estudio de las micropulsaciones magnéticas está a nuestro alcance. El magnetómetro de Baker, barato, permite captar sin dificultad esas minúsculas pulsaciones y también los efectos a veces espectaculares de las tormentas magnéticas en la ionosfera.

El ingenio en cuestión emplea un instrumento muy sensible, la balanza de torsión, que sirve para medir una fuerza utilizándola para retorcer un hilo delgado. El filamento se opone a la rotación desarrollando un par de fuerzas que va creciendo hasta que exactamente equilibra al momento generado por la fuerza aplicada. El ángulo de desviación resultante, que se halla haciendo rebotar un haz luminoso en un espejito sujeto al hilo, es proporcional a la fuerza estudiada. Con el haz procedente de un puntero de láser y un espejo del tamaño adecuado, podremos separar desviaciones de una diezmillonésima de grado.

La mayoría de las balanzas de torsión profesionales emplean fibras de cuarzo finas, resistentes e insensibles a los cambios de humedad y temperatura. Pero no es fácil hacerse con fibras de cuarzo. Baker ha descubierto que las fibras de nailon sirven perfectamente. Empecemos con un cordel de nailon multifilamentoso sedoso; cortemos un trozo de 30 cm. Desharemos ahora el torcimiento




*El haz reflejado de este puntero de láser oscila cuando la componente horizontal del campo magnético terrestre varía, siquiera sea ínfimamente, lo mismo que cuando las tormentas ionosféricas crean micropulsaciones magnéticas. Para estudiar la componente vertical del campo los imanes deben disponerse paralelos al suelo. (Nota: En esta ilustración el plano de los imanes de cero, que debe ser paralelo al plano de la moneda de un centavo, se ha rotado ligeramente para ofrecer una mejor vista del instrumento.)*

con unas pinzas para seleccionar las hebras más finas, que deben ser de unas 25 micras de grueso.

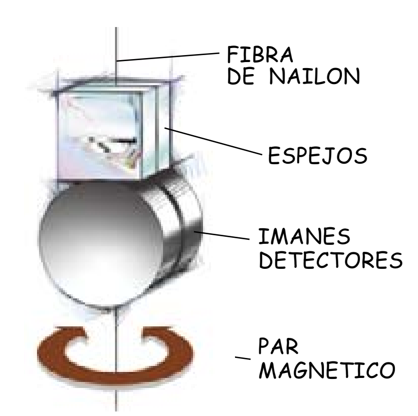
Baker instala el filamento de nailon en el interior de una sencilla montura de vidrio. Con un diamante de cristalero cortamos dos bandas de 5 cm de ancho por 15 de largo. Nos servirán de paredes verticales de la montura. Cortamos luego ocho bandas de vidrio de 1 cm de ancho por 5 de largo; con silicona los pegamos por parejas, cara con cara. Por último, pegaremos una pareja de esas banditas de vidrio al borde superior y otra pareja al borde inferior de las bandas de vidrio más largas constituidas en paredes. Las piezas pequeñas actúan como separadores entre las paredes.

Una vez seco el pegamento, se cubren los separadores horizontales de una de las paredes con una capa de cinta aislante de vinilo negra, que dé de sí; la cinta impide que el vidrio corte la fibra. Se pasa un extremo de la fibra a través del separador superior y se fija con epoxia; la aseguramos con otro trozo de cinta. La epoxia se opondrá a que la hebra resbale con el paso del tiempo. Baker genera la tracción suficiente en la fibra colgando del extremo cuatro monedas de cinco centavos durante un tiempo. Luego coloca epoxia y encinta en su posición el extremo inferior de la hebra contra el separador del borde inferior, bloqueando la tracción.

Para lograr que el nailon se retuerza reaccionando a minúsculos cambios en el campo ambiente, hay que dotar a la fibra de un imán potente. Un imán grande y masivo respondería con excesiva morosidad. Lo ideal es un imán muy intenso y ligerísimo. Tales milagros magnéticos existen bajo el nombre de imanes de tierras raras, y así se llaman porque contienen elementos de tierras raras, como el samario. Pese a sus limitadas dimensiones, albergan en su superficie unos campos magnéticos 10.000 veces más intensos que el terrestre.

En uno de esos imanes depositamos una gotita de pegamento de silicona; emparedamos el filamento entre ambos. Asegúrese de que se solapan completamente entre sí y que están perfectamente centrados respecto a la fibra al secarse el pegamento.

Baker convirtió en el reflector un espejito de tocador. Con un diamante de vidriero cortó pedacitos cuadrados de 1,5 mm de lado, que luego pegó unos con otros, reverso con reverso y centrados con la fibra, justo encima y en contacto con los imanes de tierras raras. Así el espejo y los imanes rotan como un conjunto bien equilibrado. El instrumento trabaja mejor si por superficie reflectora de la luz empleamos el reverso del espejo, pues así se evita que el haz se distorsione al atravesar el vidrio. Para eliminar la laca que cubre el revestimiento reflexivo, nos vale un bastoncillo de algodón y un poco de cetona etilmetílica. Si se prefiere no trabajar con un compuesto potencialmente tóxico, instálense los espejos del modo habitual, con la luz atravesando el vidrio.



*Trozos de espejo y una pareja de imanes potentes, sujetos a una fibra de nailon, giran ante las variaciones del entorno magnético*

En ese momento Baker pega un centavo de cobre (acuñado antes de 1982, cuando la pureza del metal era aún alta) al vidrio, detrás de los imanes. En su movimiento, éstos inducen corrientes de Foucault en el cobre, que a su vez generan su propio campo magnético, opuesto al movimiento de los imanes. El ingenioso truco de Baker amortigua enseguida las oscilaciones no deseadas y facilita la lectura del magnetómetro.

Se procede a encapsular el sensor pegando la segunda pared de vidrio con sus separadores encima de la primera. Se sellan los lados con cinta aislante negra para proteger al magnetómetro de corrientes de aire molestas. El conjunto se monta sobre una base horizontal plana. Ya tenemos la brújula de precisión. Cuando nos movamos por el entorno, los imanes deberán alinearse con el norte magnético y exhibir muy poca oscilación.

El campo terrestre confunde nuestras micropulsaciones. Conscientes de ello, empezaremos por poner a cero el instrumento antes de que éste capte señales procedentes de la ionosfera. Con ese fin, compraremos cuatro imanes tóricos, que dispondremos uno junto a otro pegados a un pequeño trozo de vidrio o madera. (Necesitaremos abrazaderas para sujetarlos en contra de su repulsión magnética mutua mientras se seca el pegamento.) El conjunto de estos imanes se pone de pie y se pega a una base autotestable de suerte tal, que el centro del conjunto se alinee con los imanes de tierras raras del sensor.

Cuando los imanes de cero se acercan al magnetómetro (unos 30 cm), el sensor empezará a bambolearse. Girará libremente en cuanto las fuerzas combinadas del campo magnético terrestre y las de los imanes tóricos casi se anulen entre sí. El período de las oscilaciones debe alargarse hasta un segundo o más, cuando todas las fuerzas se aproximen al equilibrio. Esa situación corresponde a la máxima sensibilidad del magnetómetro.

El instrumento se completa con el rayo luminoso de un puntero de láser. Sitúese el láser de modo que el haz traspase el vidrio, rebote en el espejo y alcance un muro apartado. Las fluctuaciones del campo magnético terrestre se muestran como desviaciones del haz.

Estamos listos para poner a prueba nuestro dispositivo, que acaba de mejorar Baker. Arrollando hilo conductor sobre un envoltorio de copos de avena, ha construido un par de bobinas para calibrar su invención. Recurre al dispositivo de Helmholtz formado por dos bobinas redondas (separadas una distancia igual a la mitad de su diámetro) que uniformiza el campo magnético en el centro. Las bobinas de Helmholtz pueden combinarse con fotocélulas y realimentación electrónica para inmovilizar el haz de láser cuando varíe el campo magnético terrestre. Midiendo la corriente necesaria para anular la señal, se sabe la amplitud de las fluctuaciones magnéticas. Mediante un ordenador que registre la corriente se genera un flujo de medidas que después pueden analizarse.

*Para más detalles acerca de este proyecto, consúltese el sitio Web de Roger Baker en www.eden.com/-rcbaker. Visítese el foro Web de la Society for Amateur Scientists en web2. thesphere.com/SAS/WebX.cgi. Puede también escribirse a la sociedad en 4735 Clairemont Square, Suite 179, San Diego, CA 92117, o llamar al 619-239.88.07.*

Ian Stewart

# Reparto sin envidia

En el artículo de mes pasado echamos una ojeada a algunos de los problemas matemáticos nacidos del problema de la división equitativa de una tarta; equitativa en el sentido de que, si hay *n* personas, todas queden convencidas de que su parte es por lo menos 1/*n* de la tarta. Veremos ahora cuestiones emparentadas, que han originado las partes más modernas de la teoría.

¿Dónde nos quedamos? En el caso de dos personas, el clásico algoritmo "Yo parto, tú eliges" permite un reparto equitativo. En el caso de tres o más personas existen varias posibilidades. El método de "afinamiento" permite a los sucesivos participantes ir reduciendo el tamaño de una porción presuntamente equitativa de la tarta, a condición de que si nadie más recorta esa porción la última persona en tocarla ha de aceptarla para sí. En el algoritmo de "pares sucesivos" los dos primeros dividen la tarta en partes iguales, y la tercera persona se asegura una porción de cada parte que corresponda, a su juicio, por lo menos a un tercio, negociando por separado con cada uno de los dos primeros. Y en el método de "divide y vencerás", los participantes tratan de dividir la tarta con un solo corte, de modo que alrededor de la mitad de los interesados acepten una parte equitativa de una de las porciones, mientras que las demás se sentirían satisfechas recibiendo una parte equitativa de la otra. La misma idea se aplica después a cada subtarta ya cortada, y así sucesivamente.

Los algoritmos anteriores son equitativos, pero existe un problema más sutil. Aun cuanto estén todos persuadidos de que han recibido una porción equitativa de la tarta, todavía es posible que haya quien se sienta mal tratado, por culpa de la envidia. Es posible, por ejemplo, que Tere, David y Hilario estén convencidos de que han recibido al menos un tercio del pastel; ello no impide que Tere pueda pensar que la porción de David es mayor que la suya. La porción de Tere es "justa", pero ya no se siente tan satisfecha como antes. Un reparto de la tarta tiene la propiedad de "no provocar envidia" si ninguno de los participantes considera que nadie tiene una porción mayor que la suya. Aunque un reparto sin envidia siempre será considerado equitativo, una división equitativa sí puede suscitar envidias. Por esta razón, los algoritmos de reparto sin envidias resultan más difíciles de hallar que los de mero reparto equitativo.

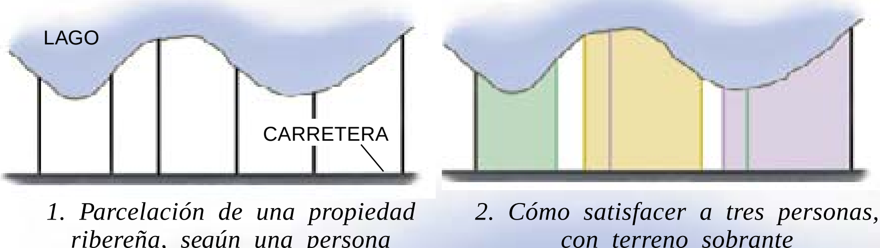
El primer algoritmo sin envidia para tres personas fue descubierto por John Selfrigde y John H. Conway a principios de los años sesenta:

**PASO 1**: Tere corta la tarta en tres porciones que tengan, a su entender, el mismo valor.

**PASO 2**: David puede, o (*a*) no hacer nada, si considera que las tres piezas son iguales o que lo son las dos mayores, o bien (*b*) recortar la que estime ser la pieza mayor, para crear empate entre dos. Se colocan aparte los recortes, que formarán "las sobras."

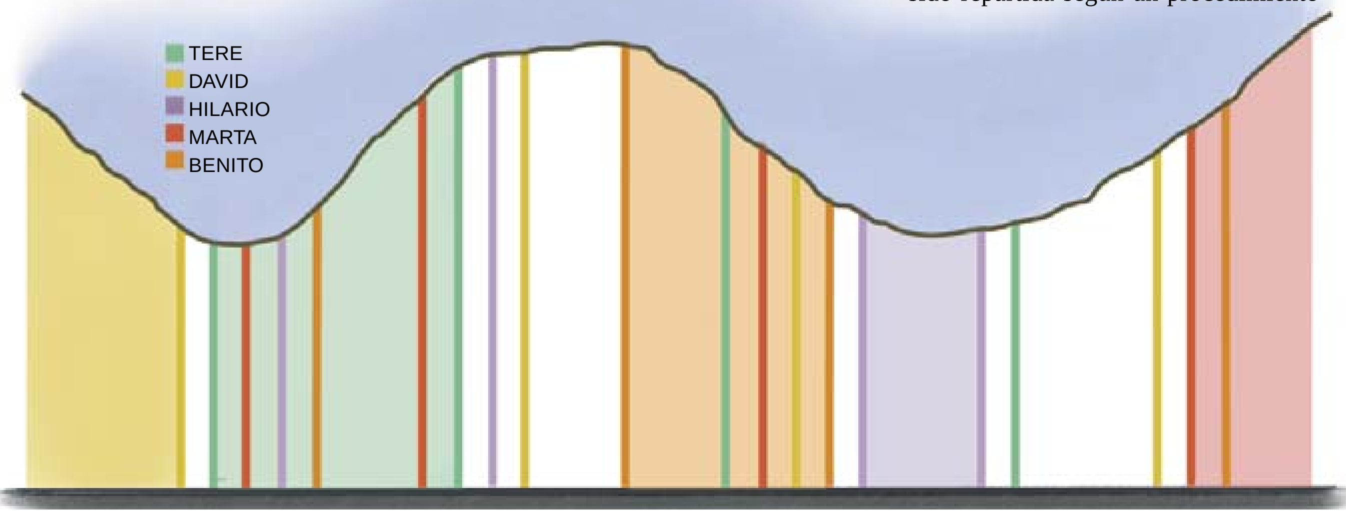
**PASO 3**: Hilario, David y Tere, en ese orden, eligen una porción: la que consideren mayor, o una de las mayores empatadas. Si en el paso 2 David ha recortado una pieza, tendrá que tomarla para sí, salvo que haya sido la elegida por Hilario.

En este estadio, parte de la tarta ha sido repartida según un procedimiento




*1. Parcelación de una propiedad ribereña, según una persona*

*2. Cómo satisfacer a tres personas, con terreno sobrante*




*3. Cómo satisfacer a cinco personas, con terreno sobrante*

en el que no puede haber envidia. Lo que falta, pues, es repartir también sin envidias las sobras.

**PASO 4**: Si David no hizo nada en el paso 2, no hay sobras, y la tarta ha quedado repartida. En otro caso, o bien David, o bien Hilario, han tomado la porción mermada. Supongamos que fue David quien optó por ella (si hubiera sido Hilario, intercámbiense de aquí en adelante los papeles de estas dos personas en la descripción del procedimiento). En tal caso, David divide las sobras en tres partes que sean, a su juicio, iguales.

**PASO 5**: Lo único que falta es que Hilario, Tere y David, en ese orden, elijan una parte de las sobras. Hilario es el primero en elegir, y no tiene motivo para sentir envidia. Tere no podrá envidiar a Hilario, cualquiera que fuere la división de los sobrantes, porque lo más que Hilario puede obtener es una parte que según Tere vale un tercio. Y no podrá envidiar a David, porque ella elige antes de que éste lo haga. Tampoco David tiene motivos de queja, pues repartió a su gusto las sobras.

En este punto, todo el mundo se quedó atascado durante 30 años. ¿Existirá un protocolo sin envidias para *n* personas? En 1995 Steven J. Brams y Alan D. Taylor descubrieron un notable protocolo para un número cualquiera de jugadores. Es francamente complicado, y no lo expondré aquí: los lectores pueden optar entre consultar el artículo de ambos "An Envy-Free Cake Division Protocol" (*American Mathematical Monthly*, enero de 1995) o el libro *Cake Cutting Algorithms*, por Jack Robertson y William Webb.

Una de las peculiaridades más interesantes de la teoría de reparto de tartas es la que Robertson y Webb llaman la "feliz discrepancia". Se podría pensar, a primera vista, que la división equitativa es más justa cuando todos están de acuerdo sobre cuánto vale cada elemento de la tarta —pues en tal caso no podrá haber disputas respecto al valor de una porción determinada. En realidad ocurre lo contrario: en cuanto los participantes disienten sobre los valores, resulta más sencillo contentarlos a todos.

Imaginemos, por ejemplo, que Tere y Hilario están aplicando el algoritmo de cortar y elegir. Tere divide la tarta en dos partes que considera del mismo valor, la mitad cada una. Si David está conforme con estas valoraciones, nada más puede hacerse. Supongamos, en cambio, que David valora las dos piezas en $^3/_5$ y $^2/_5$. Podría entonces, por alguna razón altruista, decidir conceder a Tere $^1/_{12}$ de la pieza que a su juicio es mayor (que él valora en $^1/_{20}$ de la tarta entera). Según su propia valoración, David conserva todavía $^3/_5 - ^1/_{20} = ^{11}/_{20}$ de la tarta. Para realizar su propósito, David puede dividir en 12 partes subjetivamente iguales la pieza que a su entender es mayor. Después le indica a Tere que elija la que prefiera de las doce.

Cualquiera que sea la elegida por Tere, David sigue pensando que todavía le han tocado $^{11}/_{20}$. Tere, por otra parte, tiene ante sí 12 posibles opciones, que suman en total, a su modo de ver, $^1/_2$. Por tanto, alguna de ellas vale, en su estimación, por lo menos $^1/_{24}$ de la tarta. Eligiéndola, Tere termina disponiendo de lo que considera es al menos $^{13}/_{24}$ de la tarta. Así pues, tanto Tere como Hilario están convencidos que su porción es mayor que la justa.

Lo que se intuye aquí no es que la discrepancia sobre los valores tienen que desembocar en desavenencia sobre lo que sería un reparto justo. Ello sí podría suceder, en cambio, si fuera una tercera persona la que repartiera y se empeñara en que Tere y David aceptaran una de esas porciones predeterminadas.

Tenemos otro ejemplo del mismo principio en el problema del reparto de una propiedad que mira a la playa. Supongamos que una carretera recta corra de este a oeste y que el lago haya de ser repartido entre *n* personas trazando de norte a sur líneas divisorias, de modo que cada persona reciba una parcela conexa y que todas consideren haber recibido al menos $^1/_n$ del valor total.

La solución es de una sencillez candorosa. Tomemos una fotografía aérea de la propiedad y pidámosle a cada persona que trace rectas de norte a sur, de modo que, en su estimación, la tierra quede dividida en *n* parcelas del mismo valor. Si todos dibujasen las rectas en los mismos lugares, todos quedarían satisfechos cualquiera que fuera la parcela adjudicada a cada uno. Si, por otra parte, existieran discrepancias sobre la situación de las líneas, resulta posible convencer a todos de que tienen una porción equitativa y dejar sobrante parte de la propiedad. En la figura 2 puede verse un caso típico en el que David, Tere y Hilario han aplicado un procedimiento así.

La figura 3 presenta un ejemplo más complicado, en el que Tere, David, Hilario, Marta y Benito quieren todos $^1/_5$ de la tierra. En 1969 Hugo Steinhaus demostró que lo mismo ocurre para cualesquiera elecciones de las divisorias en las que exista la más ligera discrepancia. Se puede hallar una demostración, basada en el principio de inducción matemática, en el libro de Robertson y Webb.

No siempre es posible repartir las tareas equitativamente, no, al menos, si se les imponen restricciones razonables.

## Acuse de recibo

En el artículo de junio de 1998, "La ley de promedios, derogada" mencioné que, en un paseo aleatorio bidimensional por una retícula cuadriculada infinita, la probabilidad de acabar regresando al origen es 1, mientras que en un retículo espacial cúbico tal probabilidad es menor que 1, aproximadamente, 0,35. La fuente que utilicé para este dato fue el libro *Una introducción a la teoría de probabilidades y sus aplicaciones*, vol. 1, por William Feller. Varios lectores han hecho notar que el valor indicado en el libro de Feller no es del todo correcto. David Kilbridge dice que en 1939 George N. Watson halló el valor $1 - 1/[3(18 + 12\sqrt{2} - 10\sqrt{3} - 7\sqrt{6})\,(K(2\sqrt{3} + \sqrt{6} - 2\sqrt{2} - 3))]^2$, siendo $K(z)$ igual a $^2/_\pi$ por la integral elíptica de primera especie con módulo $z^2$. (Las funciones elípticas constituyen una generalización clásica de las funciones trigonométricas, como la seno o coseno.) Su valor numérico es aproximadamente 0,340537329551.

Kilbridge dice también que la solución de mi problema sobre igualdad de las frecuencias acumuladas es aproximadamente 0,022. En el caso de dados con 2, 3, 4 o 5 lados, las correspondientes probabilidades son 1, 1, 0,222 y 0,066, respectivamente. Yuichi Tanaka, un redactor de nuestra edición japonesa, recurrió a un ordenador para calcular la probabilidad del eventual retorno al origen al pasear por una retícula hipercúbica tetradimensional. Al cabo de tres días, su programa imprimió el valor aproximado 0,193201673. (¿Existirá una fórmula como la de Watson? ¿Es algún lector experto en funciones elípticas?)

*—I.S.*

Michael Mohr

## Pianos de cola

El piano se remonta a principios del siglo xviii, cuando Bartolomeo Cristofori, constructor de clavicordios florentino, pretendiendo mejorar la expresividad de tal instrumento, concibió el "accionamiento por escape", mecanismo éste que, en forma modificada, yace en el corazón de todos los pianos modernos. A diferencia del clavicordio, caracterizado por un mecanismo de punteo directamente unido a cada tecla, el piano posee una interfaz mecánica entre las teclas y un macillo de punta cubierta de fieltro. Se trata de un complejo sistema de palancas, tornillos, muelles y cojinetes que permiten al pianista ejecutar rápidamente y producir una amplia gama de expresiones y matices.

Ese sistema de palancas amplifica el movimiento. Así, cuando la tecla recorre poco más de un centímetro desde su posición de reposo, el macillo recorre casi cuatro y medio hasta la cuerda que tiene encima. Inmediatamente antes de que el macillo choque con la cuerda, el mecanismo de accionamiento debe desacoplarse, o "escaparse", de su contacto con la tecla. Por tanto, el macillo golpea a la cuerda bajo su propio impulso y retrocede al instante, con lo que la cuerda puede vibrar libremente. Sin el efecto de escape, el macillo se atrancaría con la cuerda y se sofocaría la vibración. Los pianos modernos disponen de un accionamiento por escape doble que permite reactivar el macillo, aunque la tecla no haya sido liberada del todo; de ese modo, una nota puede ser ejecutada en repetición rápida.

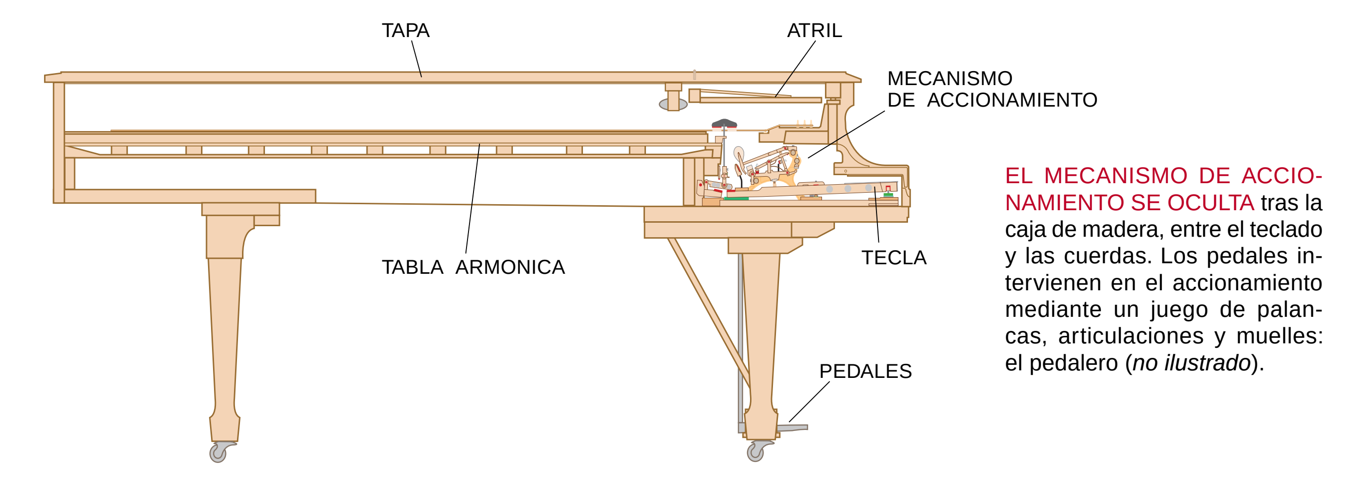
El instrumento de Cristofori disponía de un registro dinámico bastante ancho, aunque no era tan sonoro como los pianos modernos. Al paso de los años, la estructura de aquel piano fue consolidándose para admitir cuerdas más resistentes que dieran sonidos más vigorosos; se introdujeron más pedales (el de Cristofori disponía sólo del "bajo"). Aunque los detalles del piano de cola moderno difieren de un fabricante a otro, la concepción básica es la misma. En los pianos verticales, la gravedad por sí sola no devuelve el mecanismo de accionamiento a la posición de reposo. Para compensar, en los mecanismos de accionamiento verticales se incluyen unos muelles especiales, pero el tecleo difiere del peculiar de un piano de cola.

EL PUENTE transmite las vibraciones de las cuerdas cruzadas a un diafragma de madera de poco espesor, instalado debajo y llamado tabla armónica, con lo que se amplifica el sonido.

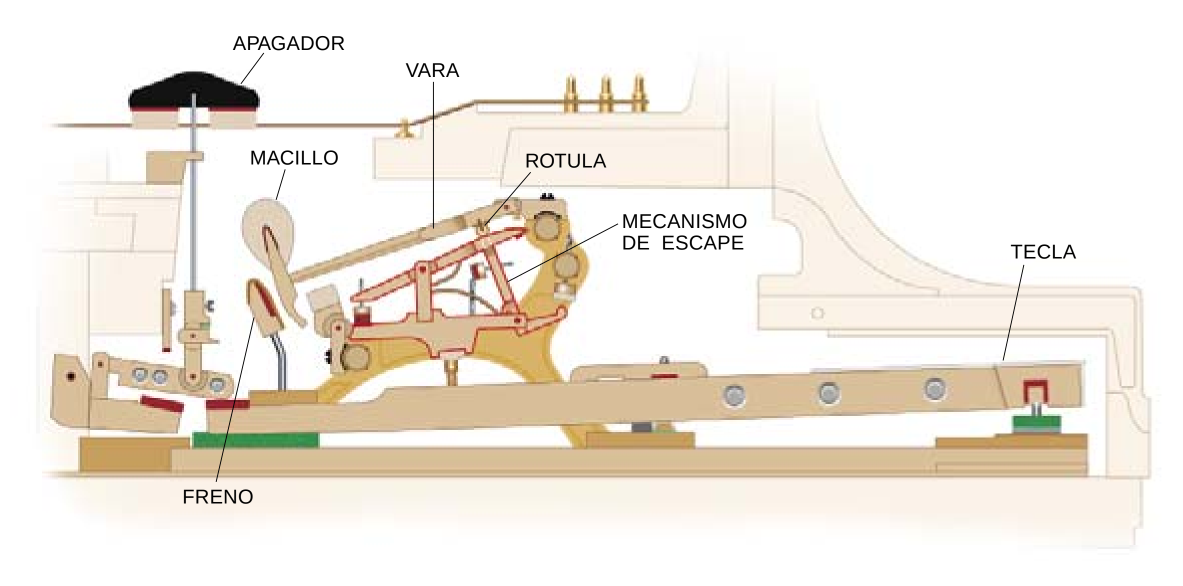
LA DISPOSICION CRUZADA DE LAS CUERDAS, o sea, con las cuerdas bajas cruzadas sobre las agudas, permite alojar las cuerdas bajas más largas, responsables de la mayor riqueza del sonido, en una caja de un tamaño dado.

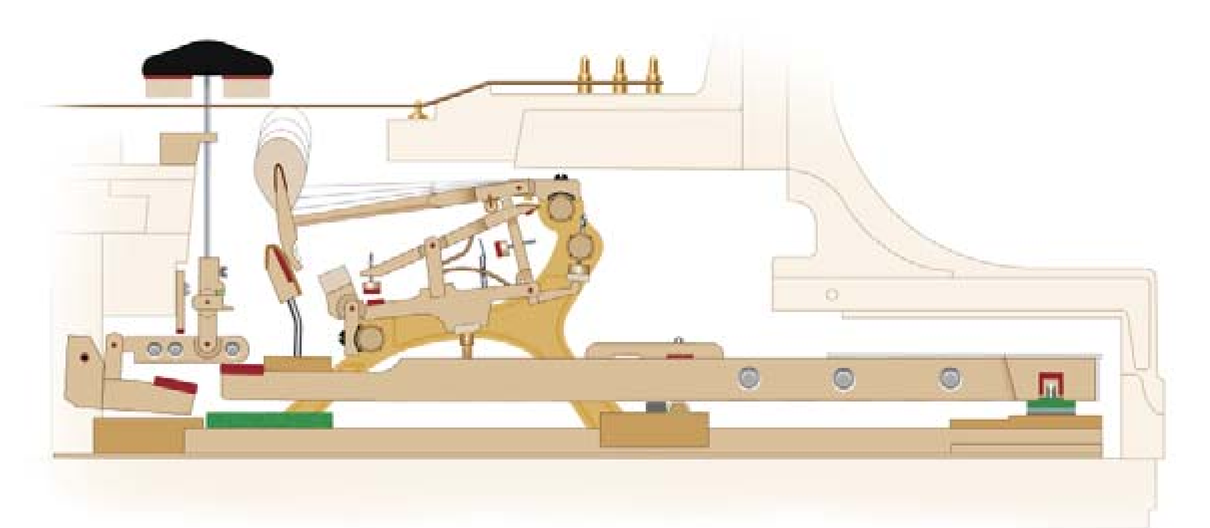
EL MARCO DE HIERRO, o arpa, está sujeto al bastidor interno del piano y mantiene a las 220 cuerdas bajo una tracción superior a las 18 toneladas.

LOS PEDALES producen efectos especiales. El pedal izquierdo atenúa el sonido desplazando todo el mecanismo de accionamiento y el teclado, de modo que los macillos agudos percutan sólo dos de las tres cuerdas por nota. El pedal izquierdo sostiene el sonido elevando a la vez todos los apagadores de las cuerdas, dejando así que éstas sigan vibrando una vez percutidas. El pedal central sostiene sólo las notas cuyos apagadores son elevados en el momento en que se acciona el pedal.




EL MECANISMO DE ACCIONAMIENTO SE OCULTA tras la caja de madera, entre el teclado y las cuerdas. Los pedales intervienen en el accionamiento mediante un juego de palancas, articulaciones y muelles: el pedalero (*no ilustrado*).



AL PULSAR UNA TECLA sube el mecanismo de escape y a su vez se eleva la vara. En dos puntos críticos del movimiento ascendente del macillo, los topes liberan elementos del mecanismo de escape de su contacto con la rótula; así, durante los últimos 1,6 mm de su movimiento, el martillo se desplaza por su propio impulso. Mientras, el apagador se ha alzado separándose de la cuerda y ésta puede vibrar al ser percutida. El macillo, después de percutir, es atrapado y retenido por el freno. Al menor aflojamiento de la presión sobre la tecla, el freno se separa y libera al macillo; se restablecen en su lugar correspondiente los componentes del mecanismo de escape. Y el mecanismo de accionamiento queda listo para recibir otra pulsación de la tecla.

# NEXOS

James Burke

## Y ahora, el tiempo

Recalé en una iglesia parisiense, el famoso *Panthéon*, para guarecerme de una tormenta llegada del Atlántico, una más. Allí filmé una secuencia sobre el tipo que nos explicaba por qué la lluvia siempre viene de esa dirección. En el Panthéon pende el gran experimento de Léon Foucault, quien en 1851 colgó una bola de cañón de 62 libras de una cuerda de piano de 67 metros. La ató a un extremo con un cordel. Liberó la bala para que se abandonara sin influencia alguna a su movimiento. El patrón de movimiento que siguió la bola en las horas inmediatas fue la primera prueba física de la teoría de Copérnico. El péndulo se balanceaba en un espacio inerte y un estilo colocado en la parte inferior de la bola trazaba una línea que se iba desplazando conforme la tierra se movía a sus pies. Lo divertido del caso es que esta demostración se convirtiera en el fundamento del pensamiento meteorológico de Buys Ballot y otros sobre la determinante influencia de la rotación oeste-este de la Tierra en los fenómenos meteorológicos. Foucault se hizo famoso de inmediato.

El nombre de Foucault resplandeció también por otras motivos. Por ejemplo, su perfeccionamiento de un regulador que mantenía los filamentos de carbono del arco eléctrico a la distancia adecuada cuando estaban incandescentes. Así se disponía de luz suficiente para iluminar lugares públicos, como los teatros donde se divertían de noche gente como el químico Henri Moissan. En su laboratorio, Moissan utilizaba el arco con fines muy distintos. Lo empleaba para alimentar un horno de arco eléctrico, en el que los filamentos de carbono alcanzaban temperaturas que prácticamente le permitieron fabricar diamantes artificiales. Moissan fue *le noodleur extraordinaire* que recibió el Nobel en 1906, aisló fluoruro sin perecer en el intento y escribió más de 300 artículos científicos. En el invierno de 1897, proporcionó a una joven polaca que vivía en París polvo de uranio, para que investigara su misteriosa capacidad para "electrificar" el aire. Se trataba del primer paso que daba Marie Sklodowska Curie en pos del descubrimiento de la naturaleza de esa "electrificación", que más tarde se llamaría radiactividad.



Marie se había casado poco antes. Había entroncado con la familia Curie, de sólida tradición igualitarista, científica y emancipada. Tenía la culpa el abuelo Paul Curie, que a principios de siglo fue seguidor de una secta, la Nueva Cristiandad, fundada por un tal Henri de Saint-Simon. Este, en un momento en que se había encontrado hambriento y en la calle, se disparó hasta seis veces en la cabeza. Sobrevivió. Quería poner al día la religión y acomodarla al mundo moderno. Sus panfletos hablaban del poder espiritual de los hombres con conocimientos prácticos y en general glorificaban el trabajo y la ética capitalista. No resulta sorprendente que, hacia 1830, ingenieros, banqueros y hombres de negocios engrosaran la Nueva Cristiandad y hablaran de amor libre (este aspecto en particular terminaría por llevar la nueva religión al fracaso).

Había entre sus seguidores alguien que hizo bueno a de Saint-Simon. Se llamaba Barthélemy-Prosper Enfantin, de profesión banquero. Erigido en sucesor, se autoarrogó en su modestia la mitad del alma de Jesús. En un viaje por Egipto en busca de esposa (que nunca encontraría) le excitó una de las ideas que rondaban también a de Saint-Simon: la construcción de canales. En 1847, el estudio de Enfantin para el Canal de Suez (una visión en la que quiso unir este y oeste) le dieron buenas razones para enfrentarse a Ferdinand, vizconde de Lesseps, que estaba construyendo el canal propuesto.

Ferdinand era un diplomático, cuyo padre, cónsul francés de Napoleón, había escogido a un soldado analfabeto para hacerlo khedive de Egipto. Con semejante jefatura, Ferdinand no podía hacer nada mal en El Cairo. Cuando propuso el Canal de Suez, el khedive halló lo que estaba buscando para conseguir que la reputación de su país subiera alguna pulgada. Ferdinand tardó diez años y empleó 25.000 obreros, pero la apertura oficial en noviembre de 1869 es uno de los mayores acontecimientos de la historia. Glitterati no es la palabra. Cena para ocho

mil, todos ellos escrupulosos con el protocolo. De los pocos que podían saltárselo era la mujer de Napoleón III de Francia, la emperatriz Eugenia, prima de Ferdinand.

Los entusiastas lectores de esta columna recordarán que una joven dama española, Eugenia, había conocido y hecho muy buenas migas con Prosper Mérimée, quien la había dulcificado. Bien, como dice la canción, hay que ser amables con las personas con que uno se encuentre. Y vaya si le salió a cuenta. Una vez Eugenia había atado el nudo imperial, Mérimée se convirtió en el autor de moda en los exquisitos salones de París, filigranas del cotilleo. También tuvo éxito en Londres, donde conoció a Antonio Panizzi, toda una institución en el Museo Británico. Si a alguien valía la pena conocer en Inglaterra era al italiano. Panizzi había llegado a Inglaterra en 1823, cuando el servicio secreto italiano se había interesado en exceso por él, tras relacionarlo con gente muy sombría, los Carbonari, que querían todo tipo de insensateces para Italia, la libertad de expresión, por ejemplo.

Las ideas libertarias de Panizzi (al igual que su acento) se modificaron de forma importante en Liverpool. Había acudido allí con una recomendación para William Roscoe. Además de gran italófilo (como yo mismo, que estoy convencido de que Italia es una enfermedad para la que afortunadamente no existe cura) y biógrafo de Lorenzo de Medici, Roscoe también era banquero, botánico, antiesclavista activista, diputado por Liverpool y editor y coleccionista de libros (y de liberales fugitivos). En 1806 escribió un clásico de las guarderías: *El Baile de la Mariposa y el Banquete del Saltamontes*, que entusiasmó de inmediato a los reyes y que publicó en John Harris, en su primera colección infantil.

Harris fue el sucesor de John Newbery, cuya imprenta londinense de St Paul's Churchyard lo convirtió en uno de los primeros editores ingleses de libros infantiles ilustrados, además de publicar a Goldsmith o a Johnson. En 1780, la lista de las ediciones de Newbery incluía los *Cuentos de la Mamá Ganso*, traducidos del francés. Su autor, Charles Perrault, había obtenido la dudosa distinción de llevar la contabilidad de la construcción de Versalles. Los hermanos Perrault eran tres: Claude diseñó una parte del Louvre y del Observatorio de París y Pierre, un recaudador de impuestos al que sorprendieron con las manos en la masa, fundó la ciencia de la hidrología y trabajó con un tipo que luego se atribuyó los hallazgos de Pierre. Edmé Mariotte parece que repitió el truco con alguno más: Huygens lo acusó de plagio.

A Mariotte podemos atribuir, casi sin duda alguna, la organización hacia 1670 de las informaciones proporcionadas por las estaciones de toda Europa, que le permitió redactar un informe sobre los vientos de Europa y del mundo. Fue el primero en plantear la teoría en cuya demostración Foucault colaboraría y que mencionábamos al inicio: que, en virtud de la rotación de la Tierra, el tiempo atmosférico se desplaza de oeste a este.

# LIBROS

## Gen

### *Naturaleza y función*

**THE STUDY OF GENE ACTION**, por Bruce Wallace y Joseph O. Falkinham III. Cornell University Press; Ithaca y Londres, 1997.

Este libro se complementa con *The Search for the Gene* ("La búsqueda del gen"), publicado en 1992 por Bruce Wallace, uno de los autores. Juntas, ambas obras constituyen una descripción concisa de los caminos seguidos por la genética, para avanzar en sus dos líneas de investigación fundamentales: *1*), qué son los genes y dónde están (primer volumen) y *2*), cómo llevan a cabo sus actividades (volumen aquí reseñado).

Se mueven en un nivel de alta divulgación. Siguiendo la pauta establecida en *The Search*, el texto de *The Study* es asequible, pero no superficial. Así, el lector interesado puede adquirir conocimientos precisos y correctos, sin demasiado esfuerzo, gracias a la didáctica elección y exposición de los temas tratados. A esto contribuyen, además, gráficos explicativos muy adecuados. Por eso los recomiendo a quien desee adquirir un conocimiento objetivo de la genética y de sus problemas.

Adapta su estructura al proceso histórico de adquisición de conocimientos sobre el modo de actuar los genes. En un primer capítulo introductorio, se presenta la problemática a abordar. Siguen otros 10 en los que, manteniendo en líneas generales un orden histórico, se pasa revista a las principales etapas del progreso en torno a la acción génica. De ese modo se nos ofrece desde la exposición de Garrod (1909) sobre los "Errores congénitos del metabolismo" (capítulo 2), donde se establece el carácter hereditario de las alteraciones del metabolismo, hasta las aplicaciones de la ingeniería genética a la acción génica (capítulo 10). Se dedica el tercero a las principales técnicas utilizadas y a los organismos que se toman como patrones de referencia.

En el "Epílogo", los autores se detienen en los peligros que puede comportar la práctica de la ingeniería genética. En algún aspecto sus temores me parecen exagerados. Por ejemplo, afirman que "por pequeño que sea el riesgo de que un nuevo ADN recombinante pueda poner en peligro el mundo natural preexistente (incluidos los seres humanos), el *accidente* es virtualmente seguro, si se generan números muy grandes de ADN recombinante." Esta conclusión quizá no sea tan objetiva como el resto del libro. Está basada en el enorme potencial que tiene la recombinación genética para generar variabilidad, según se pone de manifiesto en el modo de operar del sistema inmunitario (explicado en el capítulo 9). Por recombinación del ADN de las regiones codificadoras de este sistema, se generan secuencias que codifican anticuerpos que pueden actuar sobre prácticamente cualquier proteína extraña que pueda entrar en el cuerpo. Mas, la generación de antígenos por secuencias recombinantes de ADN, ¿es realmente un modelo aplicable a la generación de secuencias de ADN que pongan en peligro todo el mundo natural preexistente? ¿Pueden realmente existir secuencias con esta propiedad?

ANTONIO PREVOSTI

## Matemática

### *Números y razones*

**ROGER BACON AND THE SCIENCES. COMMEMORATIVE ESSAYS**. Dirigido por Jeremiah Hackett. Brill; Leiden, 1997. **ECONOMY AND NATURE IN THE FOURTEENTH CENTURY**, por Joel Kaye. Cambridge University Press; Cambridge, 1998. **PEDRO NUNES (1502-1578). HIS LOST ALGEBRA AND OTHER DISCOVERIES**. Edición preparada por John R.C. Martyn. Peter Lang; Nueva York, 1996.

**PI: A SOURCE BOOK**. Compilación de Lennart Berggren, Jonathan Borwein y Peter Borwein. Springer; Heidelberg, 1996. **AN IMAGINARY TALE. THE STORY OF $\sqrt{-1}$**, por Paul J. Nahin. Princeton University Press; Princeton, 1998.

**GEORGE BOOLE. SELECTED MANUSCRIPTS ON LOGIC AND ITS PHILOSOPHY**. Preparado por Ivor Grattan-Guinness y Gérard Bornet. Birkhäuser Verlag; Basilea, 1997. **A LOGICAL JOURNEY: FROM GÖDEL TO PHILOSOPHY**, por Hao Wang. The MIT Press; Cambridge, 1996. **TRUTH IN MATHEMATICS**. Dirigido por H. G. Dales y G. Oliveri. Clarendon Press-Oxford University Press; Oxford, 1998.

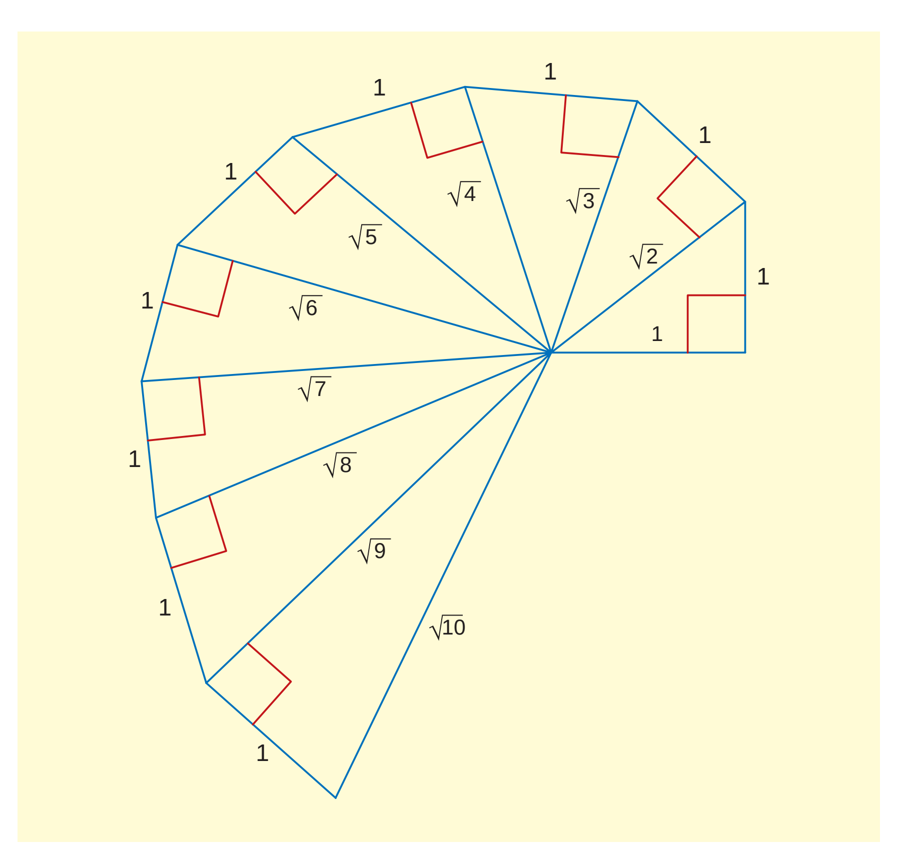
En la configuración del pensamiento matemático occidental ha habido momentos y temas que, por su interés, merecen rescatarse del académico confinamiento en revistas minoritarias donde dormitan. Uno de esos períodos cubre nada menos que cuatro siglos, de la centuria decimotercera a las vísperas galileanas. Es el arco barrido por los tres primeros libros.

*Roger Bacon and the Sciences. Commemorative Essays* compendia, a través del franciscano, el saber exacto en la Europa bajomedieval. Convergían, en Oxford, donde enseñó, tradiciones muy dispares, revitalizadas con las traducciones de textos clásicos y medievales que asimiló su maestro Robert Grosseteste. No escribió Roger Bacon (1214-1290) sobre matemática en la forma sistemática en que lo hizo sobre la luz, la visión o la fuerza, pero expuso la necesidad de abordarlas con óptica matemática y experimental.

Defiende la aritmética y geometría de las artes liberales contra los abusos de numerólogos. Escribió un práctico *Communia mathematica* y un tratado de geometría, anexo y truncado. Conoce los *Elementos* de Euclides, las esferas según Teodosio y los principios arquitectónicos de Vitruvio. Emplea la *Optica* de Ptolomeo para la geometría, con inclusiones de Arquímedes y Apolonio. Boecio y Alhazén ejercen en él una influencia determinante. Confiesa haber aprendido mucho de sus contemporáneos, "el maestro Juan de Londres y el maestro Pedro de Maricourt"; al primero se le asocia con *De curvis superficiebus*, texto de geometría; del segundo sólo nos ha llegado una carta sobre el magnetismo y un trabajo sobre construcción de astrolabios.

La matemática bajomedieval giraba en torno a las proporciones y sus permutaciones. A diferencia de Euclides, Bacon se mostró partidario de apoyar en un criterio aritmético la igualdad entre razones. En el siglo siguiente las soluciones partirán de las mesas de cambio. Tal es la novedad que nos trae Kaye en *Economy and Nature in the Fourteenth Century*. De acuerdo con su tesis, la filosofía natural vino determinada por el proceso social y económico de monetarización que transformó la sociedad europea.

Siguiendo la *Etica* de Aristóteles, las transacciones pertenecían al dominio de la moral. La idea de la justicia se ilumina con el *De arithmetica* de Boecio para los escolásticos. Justo es lo proporcionado. Los dos tipos de justicia, distributiva y directiva, se ordenan a la consecución del justo medio. La primera se imparte en proporción al trabajo y méritos; se trata de una equivalencia "geométrica". La justicia directiva se aplica a los casos en que se supone que los sujetos merecen idéntica recompensa. El dinero del intercambio o satisfacción de una deuda operaba como un *continuum* de valor, que medía y numeraba.




*Espiral de triángulos de Theodorus*

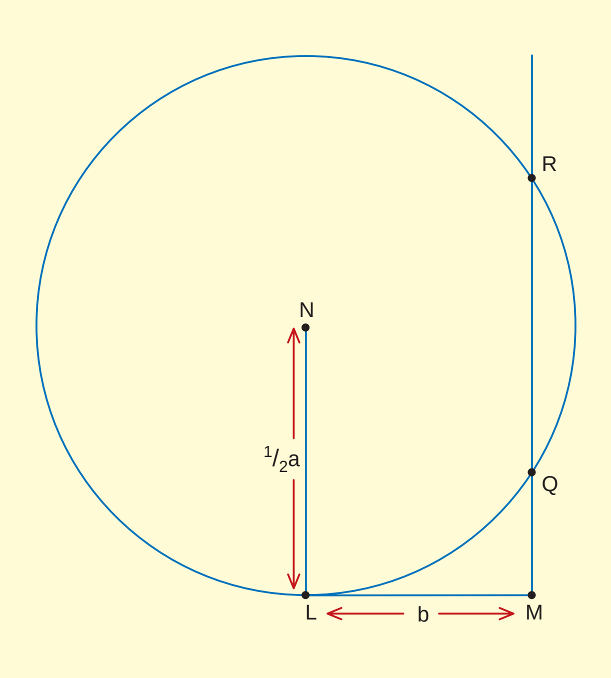
Los maestros del Colegio Merton de Oxford avanzan un paso más y someten todo tipo de saber a medición. Aplican reglas matemáticas y esquemas cuantitativos (experimentados en transacciones mercantiles) a cualidades y movimientos, a sus comienzos y términos, al intervalo continuo entre ambos extremos, lo que en jerga se llama "latitud de formas". Thomas Bradwardine, uno de los calculadores de mayor peso, introduce la primera función matemática de la medición del movimiento en *De proportione velocitatum in motibus*. Bradwardine había sido ecónomo de la universidad, es decir, administrador de sus propiedades y fincas. También recayó ese cargo en William Heytesbury, lógico y filósofo de la naturaleza. Los registros de Heytesbury nos hablan de dos semanas de trabajo repasando cuentas, revisando cosechas y recaudando en moneda o en especie lo debido. Por los días en que llevaba empeñado en ese trabajo terminó las *Regulae solvendi sophismata*, catecismo de la nueva física.

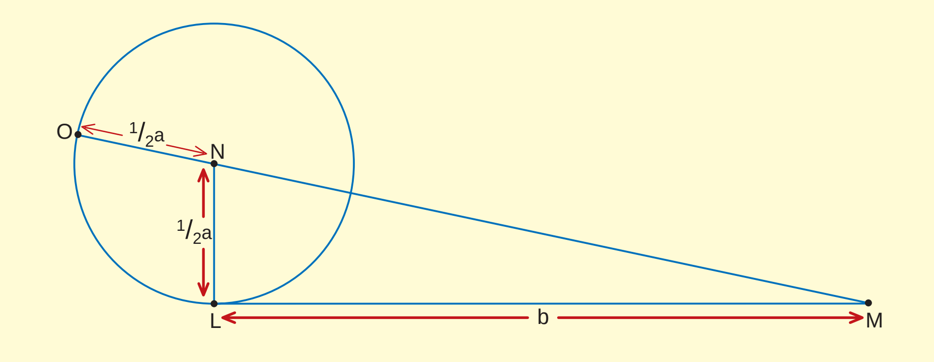
La propia relación económica sufre transformaciones. En el caso del préstamo o *mutuum*, la igualdad ideal se determinaba aritméticamente: la suma prestada señalaba el montante a devolver. Pero empezaron a cuestionarse las exigencias tradicionales de igualdad aritmética al reflexionar sobre la justicia en la compra (*emptio*) y venta (*venditio*). Peter John Olivi reconoce el papel de la *aestimatio* y la *probabilitas* en la fijación del precio. Ante el viejo problema jurídico de si debía compensarse al prestamista por el *lucrum cessans*, entiende que ha de indemnizarse con lo equivalente al beneficio probable. John Duns Scotus concebía la igualdad dentro de una gran amplitud *(magna latitudo*) de equivalencia en el intercambio.

De Oxford a París. Las *Ordonnances* reales ponen de relieve los continuos cambios y depreciaciones de la moneda por culpa de guerras, hambrunas y peste. En momentos de rápida devaluación, la sustitución en las mesas de los cambistas se hacía varias veces al año. Entre 1337 y 1360, hubo hasta 85 modificaciones. Las *Ordonnances* listaban las equivalencias proporcionales entre monedas viejas y nuevas (a veces había hasta 12 monedas implicadas), de suerte que las rentas, deudas y contratos pudieran pagarse en valores proporcionales.

De París a Salamanca, donde, un siglo más tarde, se forma Nunes, autor de una de las primeras álgebras modernas ensayada en un opúsculo *ad usum delphini* que John R. C. Martyn ha exhumado en *Pedro Nunes (1502-1578). His lost algebra and other discoveries*. Médico de la familia real, preceptor de los infantes y cosmógrafo, enseñó retórica, lógica y metafísica en la Universidad de Lisboa. Para él se creó la cátedra de matemática en la de Coimbra. Impulsó aquí el estudio del álgebra y, siendo experto cartógrafo, acudía regularmente a Lisboa para explicar a los pilotos el manejo correcto de los instrumentos náuticos, así como la interpretación de derrotas y cartas marinas. De los escritos recuperados en Evora deduce Martyn la asociación entre el círculo palatino, integrado por Nicolas Clenardo, Jerónimo Cardoso, André de Resende, Jorge Coelho, António Pinheiro y Francisco de Holanda con Nunes al frente, y la ciencia renacentista, en unos momentos en que el rey Juan III y su cuñado Carlos V controlaban el mundo conocido.

El grueso de esta obra se dedica a unas páginas escritas en torno a 1533, que habían permanecido inéditas. Nunes las redactó para enseñar álgebra al heredero de la corona y a otros jóvenes de la nobleza de Evora. Tras compararlo con su *Libro de Algebra en Arithmetica y Geometria*,




*Construcciones geométricas de Descartes de ambas raíces positivas; a la izquierda, en* $z^2 = az + b^2$*, y, arriba, en* $z^2 = az - b^2$*. Siendo* a *y* $b^2$ *positivos*

publicado en 1567 en Amberes, resalta la identidad de plan. El manuscrito de 1533 avanza la estructura que se mantendrá y desarrollará en la edición impresa. El sumario estaba escrito en portugués; el libro en castellano. Aquél consta de 24 páginas, éste de unas 600. Los ejemplos, problemas y aclaraciones extienden un esquema pergeñado. Arranca el libro con una larga introducción donde establece las reglas generales de las conjugaciones simples y compuestas, reservadas en el sumario a un tratamiento posterior. Mas a partir de ahí la equivalencia es plena: suma, resta, multiplicación, división, quebrados, raíces, proporciones, igualación y reglas del álgebra.

El interés de ciertos números trasciende las fronteras de la periodización histórica. Acontece así con *e* o la $\sqrt{-1}$. Con π. Desde el amanecer de la matemática hasta hoy se suceden las incursiones sobre el valor exacto y el significado de la razón de la circunferencia de un círculo cualquiera al diámetro del mismo. Seguir su curso milenario es seguir la historia de la geometría, el análisis y funciones, álgebra y teoría de números. Lo vemos en la extensa selección intitulada *Pi: A source book* y en *An imaginary tale*, que incluye el artículo que *Investigación y Ciencia* dedicó a Ramanujan, erróneamente asignado aquí a "Science and Application".

A grandes rasgos, se observa en el estudio de π un período inicial, geométrico, que llega hasta Newton. Se busca una aproximación a la razón entre circunferencia y diámetro mediante el cálculo de perímetros o áreas de polígonos regulares inscritos y circunscritos.

Leemos un primer esbozo en el papiro matemático de Rhind (c.1650 a.C.) y un tratamiento magistral en *Sobre medición del círculo* de Arquímedes de Siracusa, que aborda el valor de pi mediante el cálculo de las áreas de polígonos equiláteros con un número de lados creciente. Por técnicas arquimedeanas se irán conociendo hasta 35 decimales.

El segundo período duró más de cien años. Se emplean ahora expresiones analíticas. El descubrimiento por Gregory de las series arcotangentes condujo a algoritmos convergentes. Con las series de Gregory obtenía Machin en 1706 un valor de π de cien decimales. El tercer período, que se extiende hasta finales del siglo XIX, investigó el estatuto matemático de π. Lambert demostró en 1761 la irracionalidad de ese número; Lindemann, su trascendencia en 1882. En el siglo XX se ha dado un paso adelante para saber si π, no obstante su irracionalidad y transcendentalidad, es normal. En 1976 Brent y Salamin idearon, cada uno por su lado, un algoritmo de aproximación fundado en integrales elípticas que producían convergencia cuadrática en π. No ha sido el último.

Para la mayoría de las aplicaciones prácticas, basta un valor de pi que sea correcto hasta los primeros 10 decimales. Pero a veces se requiere mayor precisión en situaciones en que intervienen otras constantes y funciones elementales. Con la llegada de los computadores la cifra se ha multiplicado astronómicamente desde los discretos 2037 obtenidos en el ENIAC, en 1949.

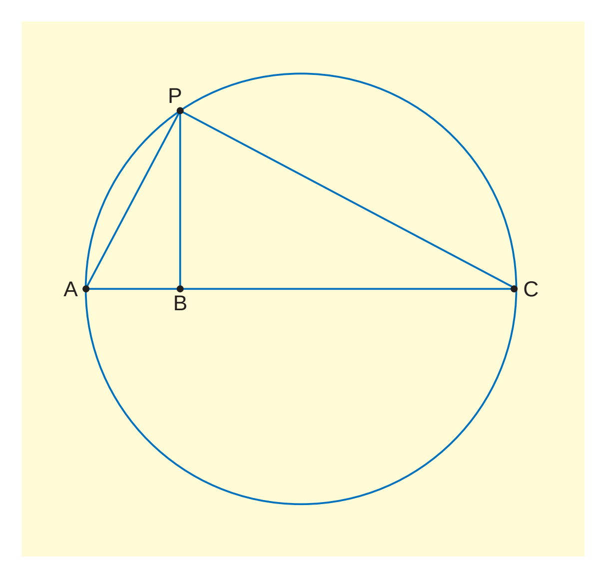
De historia bimilenaria también, los números imaginarios comparten con π su presencia ubicua en la ciencia y la técnica. Los encontramos en hidrodinámica y teoría eléctrica. El sistema informático del vuelo de la lanzadera espacial se sirve de ellos para su navegación; los químicos de proteínas, para la manipulación tridimensional de los modelos. De la necesidad de contar su desarrollo hablaban ya a mediados de siglo pasado William Rowan Hamilton y Augustus de Morgan. Le escribía aquél a éste un 13 de enero de 1852: "Usted o yo deberíamos estudiar la historia de $\sqrt{-1}$. Yo creo que usted." A los pocos días respondía Morgan: "En lo relativo a la historia de $\sqrt{-1}$, no sería trabajo pequeño recontarla desde los hindúes hasta nuestros días." Ni Hamilton ni Morgan la escribieron. El primer empeño serio, *An imaginary tale, the story of* $\sqrt{-1}$, es mérito de Paul J. Nahin.

Suele introducirse en las aulas el número imaginario $\sqrt{-1}$ con explicaciones de la siguiente guisa: "La ecuación real $x^2 + 1 = 0$ condujo a la invención de *i* (y también de –*i*). Esa es la solución." Así de fácil, así de falso. Cuando los primeros matemáticos abordaron la expresión $x^2 + 1 = 0$ y otras cuadráticas cerraron simplemente los ojos y las consideraron "imposibles". El interés por $\sqrt{-1}$ no vino de las ecuaciones cuadráticas, sino de las cúbicas que tenían soluciones reales y para las cuales la fórmula de Cardano producía respuestas con componentes imaginarios.

Como en π, aquí hemos de remontarnos a Egipto, al problema 14 del Papiro Matemático de Moscú.

Se reseña cómo calcular el volumen ($V$) de una pirámide cuadrada truncada, el *frustum* de la pirámide. De su exposición se desprende que los egipcios conocían la fórmula $V = 1/3\, h\ (a^2 + ab + b^2)$, donde $a$ y $b$ son las longitudes de los lados de los cuadrados inferior y superior, respectivamente, y $h$ la altura. A esta fórmula se le ha llamado la pieza maestra de la geometría egipcia. El desarrollo exigía el recurso a la raíz par de un número negativo. (Debemos a Leonhard Euler el símbolo $i$ para denotar la raíz cuadrada de –1.)

En un pasaje de la *Estereometría* de Herón de Alejandría volvemos a encontrar la aparición de una raíz cuadrada de un número negativo. Pero las ecuaciones que daban por solución un número negativo, insistía Diofanto, y cuantos algebristas le secundaron, eran "absurdas" y, por ende, rechazables. Descartes introdujo en *La Geometrie* el término *imaginario* para tales números, pero seguía asociando lo imaginario a la imposibilidad de una construcción geométrica.



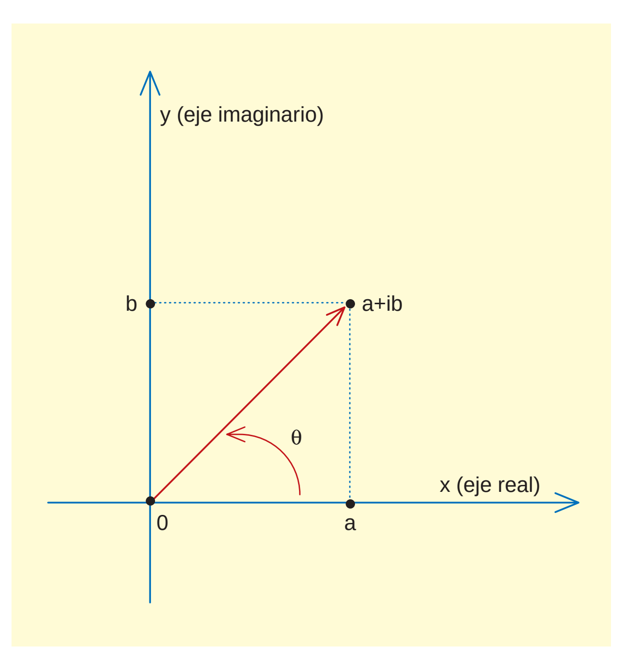

*Construcción de Wallis de la media proporcional*
$[BP = \sqrt{(AB)(BC)}]$

La respuesta no aquietó a John Wallis, quien sí se preguntó por el significado físico de $\sqrt{-1}$. En *Arithmetica Infinitorum*, Wallis argumentaba de la forma siguiente: Puesto que $a/0$, con $a$ mayor que 0, es infinitud positiva, y puesto que $a/b$, con $b$ menor que 0, es un número negativo, entonces este número negativo debe ser *mayor* que la infinitud positiva, porque el denominador en el segundo caso es menor que el denominador en el primer caso (es decir, $b$ menor que 0). Así llegó a la conclusión aporética de que un número negativo era a la vez menor que cero y mayor que el infinito positivo.

Más de cien años después del intento fallido de Wallis por domeñar geométricamente los números complejos, el problema fue resuelto por Caspar Wessel. Lo hizo en un contexto cartográfico. Presentó su trabajo, "Sobre la representación analítica de la dirección: un ensayo", a la Real Academia Danesa de Ciencias, en cuyas *Memorias* apareció en 1799. Al estar escrito en danés, pasó inadvertido. No se redescubrió hasta 1895. Para Wessel, como para nosotros, un número complejo es el punto $a + ib$ en el plano complejo (con $a$ y $b$ números reales) o el radio vector desde el origen hasta ese punto. La introducción del plano complejo en matemática por Wessel expandió el concepto de *número*. Antes de Wessel todos los números conocidos eran reales, limitados al eje $x$ unidimensional, o eje real. Tras Wessel, el dominio de los números posibles se expandió al plano bidimensional, infinito en todas las direcciones.

Aunque la brillante contribución de Wessel no se patentizó hasta mucho más tarde, sus ideas estaban en el ambiente. En 1806 aparecieron dos escritos que desarrollaban el plano complejo de Wessel y su asociación del eje vertical con el eje de los imaginarios, a saber, el *Ensayo sobre la interpretación geométrica de las cantidades imaginarias* —donde se introducía la noción de módulo—, de Jean-Robert Argand, y el artículo que Adrien-Quentin Buée publicó en las *Philosophical Transactions* de Londres, en que señala con nitidez que "$\sqrt{-1}$ es el signo de perpendicularidad." En 1835 Hamilton presentó a la Academia de Irlanda una ponencia titulada "Teoría de las funciones conjugadas o pares algebraicos: con un ensayo preliminar sobre el álgebra como ciencia del tiempo puro". Definió los pares ordenados de los números reales. Su par "puro y abstracto" $(a,b)$ es otra manera de escribir $a + ib$. Por el tiempo en que Hamilton publicó su obra sobre los pares, la interpretación geométrica había recibido ya el beneplácito de Carl Friedrich Gauss.

Pero a mediados del siglo había empezado un proceso de reinterpretación de la matemática y sus límites con la filosofía. Ivor Grattan-Guinness y Gérard Bornet lo explicitan en *George Boole. Selected manuscripts on Logic and its Philosophy*. Nacido en 1815 y autodidacta, Boole enseñó matemática en el Queen's College en Cork. Desarrolló un álgebra que podía emplearse para manipular funciones lógicas abstractas, convencido de que "la matemática trata de operaciones consideradas en sí mismas, independientemente de los objetos a los que puede aplicarse".




*Geometría de Wessel sobre números complejos*

Expuso su pensamiento en *A mathematical analysis of Logic*, que apareció el mismo día de 1847 que *Formal Logic* de Augustus De Morgan. Tras un período de reflexión de siete años publicó un segundo libro, de mayor peso, *An investigation of the laws of thought*. Dominó el cálculo diferencial e integral, sobre los que nos legó óptimos manuales. Dejó muchos manuscritos, sobre todo, para un libro



*George Boole (1815-1864)*

en torno a “The Philosophy of Logic”.

Boole articuló la lógica extensional, o lógica de clases, donde los conjuntos aparecían denotados por $x$, $y$, $z$, …, mientras que los símbolos $X$, $Y$, $Z$, …, representaban miembros individuales. La clase universal se designaba por 1; la clase vacía, por 0. Con el par $xy$ simbolizaba la intersección de dos conjuntos; empleó $x + y$ para denotar el conjunto formado por todos los elementos de $x$ y de $y$. El complementario de $x$ era $1–x$. Para indicar la relación de inclusión ($x$ está contenido en $y$), escribía $xy = x$.

Boole creía que la mente nos garantiza ciertos procesos elementales de razonamiento, los axiomas de la lógica. Por ejemplo, la ley de contradicción y la propiedad conmutativa de la intersección. A partir de esos y otros axiomas, se propuso deducir las leyes del razonamiento. De entrada aceptó que el cálculo de clases podía interpretarse como un cálculo de proposiciones. En otro orden su preocupación principal tuvo que ver con la versión algebraica del cálculo fundada en las series de Taylor, propuesta por J. L. Lagrange.

La línea iniciada por Boole culmina en Kurt Friedrich Gödel (1906-1978). En medio, David Hilbert y su escuela demostraron la consistencia de los sistemas formales y creyeron hallarse ante la demostración de la consistencia de la aritmética y la consistencia de la teoría de conjuntos. En plena euforia aparece “Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I”, donde Gödel prueba que no puede establecerse tal consistencia de la aritmética y de la teoría de conjuntos. Resultado que él derivaba del teorema de incompletitud, según el cual si una teoría formal $T$, adecuada para abarcar la teoría de números, es consistente y si los axiomas de la teoría formal de la aritmética son axiomas o teoremas de $T$, entonces $T$ es incompleta.

Nadie ha conocido mejor a Gödel que Hao Wang, autor de un póstumo *A Logical Journey*. El que ha sido reconocido por segundo Aristóteles, el mayor lógico del siglo XX o, como diría Willard Van Orman Quine, “el descubridor de la verdad matemática más significativa de este siglo”, debe tal fama en buena medida al “teorema” que lleva su nombre y ocupa un lugar central en el desarrollo de la teoría de la recursión y la teoría de la prueba. Wang nos presenta sus múltiples implicaciones: la matemática es inagotable, cualquier teoría formal coherente (consistente) de la misma debe contener proposiciones indecidibles, no hay computador capaz de demostrar todas las proposiciones verdaderas de la matemática, ningún sistema de la matemática puede ser completo, etcétera.

En 1924 Gödel entró en la Universidad de Viena para estudiar física teórica. Buscando mayor precisión, pasó a exactas en 1926 para sumergirse, dos años más tarde, en lógica matemática. Miembro asiduo de las reuniones promovidas por su maestro Moritz Schlick, coincidía en el “Círculo de Viena” con Hans Hahn, Rudolf Carnap y Karl Menger. Cerró esa etapa estudiantil en 1929 con la tesis doctoral donde demostraba la completitud de la lógica de predicados.

Mantenía el Círculo que en la lógica matemática, herramienta del análisis filosófico, se hallaba la clave para entender los fundamentos de la matemática y crear un nuevo empirismo o positivismo. Entre sus proclamas, la de que las verdades matemáticas eran analíticas, tautológicas, sin contenido, como los enunciados de que “tres pies son una yarda” por usar el ejemplo de Russell. “He sido desde 1925 un realista conceptual y matemático”, responderá Gödel en su rechazo del empirismo lógico del Círculo. Pese a haber respirado su atmósfera, es categórico en su oposición al logicismo. “Nunca sostuve la opinión de que la matemática no sea otra cosa que una sintaxis del lenguaje. De los resultados de mi labor se deduce lo contrario.” Repudia la concepción sintáctica porque la intuición matemática no puede sustituirse por convenciones sobre el uso de símbolos y sus aplicaciones, porque las proposiciones matemáticas no están vacías de contenido y porque la validez de la matemática no es compatible con un empirismo estricto.

Ese es el problema de fondo que se vuelve a plantear en *Truth in Mathematics*, volumen que recoge las ponencias sobre la cuestión de la verdad en una conferencia celebrada en Mussomeli hace cuatro años. Predomina la postura forma-

lista, que asocia la matemática a la ciencia de la prueba. Partimos de axiomas aceptados, esperamos que tales axiomas no sean contradictorios y deducimos a partir de los axiomas mediante un sistema lógico de primer orden. La interpretación que demos a los axiomas es secundaria. Sólo interesa la validez de las deducciones. Los resultados obtenidos así de los axiomas son teoremas, y, por ende, verdaderos.

Los orígenes del debate sobre la naturaleza de la verdad en matemática se remontan a la escuela de Pitágoras en el siglo V a.C. Descubrió que —dicho en nuestros términos—no existe ningún número racional $x$ tal que $x^2 = 2$. De ello se sigue que $\sqrt{2}$ no es un múltiplo racional de 1 y que hay segmentos (de longitud $\sqrt{2}$ y 1) que son inconmensurables. El descubrimiento cuestionaba la identificación entre número y geometría. ¿Cuál es el concepto "verdadero" de número? ¿Existe realmente $\sqrt{2}$ ?

Las especulaciones contemporáneas relativas a la verdad de las expresiones matemáticas parten de Kant y su tesis de que las expresiones matemáticas son *juicios sintéticos a priori*. Para el filósofo de Könisberg, un juicio *analítico* es un juicio expresado por una sentencia en que el predicado no aumenta la información encerrada en el concepto del sujeto. Un juicio *sintético* es un juicio expresado por un enunciado en que el predicado incrementa la información contenida en el concepto del sujeto, por ejemplo, $\sqrt{2}$ es un número irracional. Decir que $\sqrt{2}$ es un número real positivo $m$ tal que $m^2 = 2$ no da inmediatamente ninguna indicación de si $m$ es racional o irracional. Por consiguiente, aseverar que $\sqrt{2}$ es un número irracional es un modo de extender la información contenida a través del simple análisis de la definición de $\sqrt{2}$ . Además, un juicio es verdadero *a priori* si es verdadero independientemente de la experiencia. Por último, un juicio es verdadero *a posteriori* si sólo la experiencia puede justificar la atribución de verdad al mismo.

En la filosofía kantiana, los juicios matemáticos son sintéticos *a priori* porque son verdaderos independientemente de la experiencia y, al propio tiempo, extienden el conocimiento que podemos obtener de la simple enumeración de las propiedades del sujeto. Los conceptos matemáticos deben ser *a priori* y, por tanto, no pueden *abstraerse* de la experiencia en razón de la necesidad y universalidad de los juicios obtenidos a partir de su aplicación; si los conceptos matemáticos fueran empíricos, los juicios en que ellos se encontraran tendrían siempre la contingencia típica de juicios que sólo pueden justificarse por inducción.

En la onda kantiana, se producen varios movimientos. Así el intuicionismo de L.E.J. Brouwer, según el cual los objetos matemáticos parten de la forma a priori de actos conscientes que generan objetos empíricos. A diferencia de Kant, sin embargo, Brouwer rechazaba la aprioridad del espacio y fundaba las matemáticas exclusivamente en una concepción refinada de la intuición del tiempo. En el constructivismo de Errett Bishop, se aceptan que los objetos matemáticos sean entidades abstractas. Expulsa de la lógica la ley del tercio excluso y de la teoría de conjuntos el axioma pleno de elección.

LUIS ALONSO

# Ecología

## *Molecular*

**MOLECULAR APPROACHES TO ECOLOGY AND EVOLUTION**. VV.AA. Organizado y editado por Rob DeSalle y Bernd Schierwater. Birkhäuser; Boston, 1998.

Este libro puede ser útil para cuantos nos sentimos interesados por la evolución de los organismos asociada a su ocupación de entornos idóneos. Nos ayuda a actualizar el conocimiento de los mecanismos fundamentales de adquisición, expresión y propagación de características heredables.

Existe un ininterrumpido diálogo, que no excluye el conflicto, entre crear, introducir y conservar un mensaje modificable, expresado por una secuencia definitoria, y el grado de aceptación del resultado de ese proceso en las condiciones de entorno. Así se reformula la selección natural darwinista haciendo sitio para una prodigiosa maquinaria genética, que se va revelando por una técnica humana, cuya evolución no es menos admirable. Los experimentos y análisis avanzados se limitan a unas pocas estirpes preferidas, pero en buena medida serán generalizables. La selección natural, en cada ámbito de acción, operará como el distanciamiento que aparece y va aumentando entre los corredores de una maratón.

El número de especialistas que intervienen en el libro (33) augura una buena cobertura de las cuestiones suscitadas, aunque quizás a expensas de la coherencia y de la deseable unidad de estilo. La presente noticia se limitará a comentarios, ingenuos, de un naturalista a la antigua con voluntad de reciclarse.

La base material de la vida se organiza sobre la marcha, en un proceso histórico que encuentra expresión en un sistema informáticamente unificado e irreversible. Los aspectos considerados en el libro son muy diversos; diríase que los rápidos avances en áreas de desarrollo técnico dispar han dificultado una síntesis más significativa y accesible.

En su abundante información, se recuerdan notables homologías entre órganos diversos y complicados, aunque puede sorprender la falta de datos en lo que concierne a la regulación del número de células y su expresión en la morfogénesis. Apenas se hace referencia a animales modulares y a plantas vasculares, con los problemas peculiares que plantean. Los genes canonizados como zootípicos son los del desarrollo ("homeobox-containing"). Pasan la prueba histórica de su duplicación que permite una posible diferenciación. Son significativos los procesos de involución de la información, conseguida al combinar varios determinantes que acaban funcionando como uno solo. Esta integración debe contribuir a hacer irreversibles los pasos. Las esponjas contienen genes homeobox comparables con los de metazoos superiores.

Es natural que los autores no vean con simpatía el concepto de arquetipo ni otras formas de un pensamiento quizá trasnochado. La interpretación física y filosófica de las sucesivas decisiones es difícil de percibir en su conjunto, o no vale la pena hacerlo, pienso yo, fuera de algún concepto muy generalizable, a la vez que sencillo y profundo, de mínima acción física. Una cosa son los organismos viables y otra las trayectorias resultantes de los "experimentos" que han conducido a ellos sin obstáculo y que se expresan por un sistema de ramificaciones en el tiempo. Se asigna exagerada importancia a la configuración final de los organismos viables, en cuya generación se han sucedido numerosos pasos aleatorios.

Los problemas técnicos ocupan amplio espacio. El naturalista ha de saber que la información que está interpretando queda limitada y configurada por la técnica, más que por la lógica que instó la exploración. Hay que ir construyendo la ciencia con el material de que se dispone, cuyo acervo aumenta en progresión geométrica. La interpretación de las secuencias lineales de los sistemas genéticos se hace difícil por la existencia de intercalaciones no funcionales y de repeticiones que parece que no aportan información nueva, aunque esto sea difícil de asegurar. Por lo menos pueden tener interés histórico, testigos de acontecimientos del pasado.

Sin salirse de la selección natural, los procedimientos de la evolución han cambiado a lo largo de la historia. Una cosa son los organismos que funcionan y viven y otra las relaciones de descendencia resultado de los procesos que han generado y conducido a la evolución. Es natural que las afinidades en el genoma puedan asociarse con relaciones genéticas, morfológicas y ecológicas. Y existen paralelismos o analogías en las interacciones entre especies a lo largo de la sucesión ecológica, o en las presiones del entorno y sus cambios, que pueden haber constituido escenarios donde se han desarrollado procesos evolutivos parangonables.



*Chimpancé* (Pan troglodytes) *en su entorno de la selva de Gabón*

DeSalle y Schierwater examinan la genética y su contribución directa al desciframiento del origen de las especies. Quizás empecemos a reconocer la diferencia que hay entre la genética que conduce a la formación de especies y la genética que resulta de que existan especies. La genética incluye también la autoorganización de sus propios mecanismos. La formación prezigótica o postzigótica de barreras levantadas contra el cambio de genes puede ser importante. El estudio de pequeños grupos de especies congéneres en bacterias, *Crocodylus*, *Cicindela* o *Pinus*, permite comprobar diferencias genéticas y características "heterodoxas", que podrían ser propias de plantas vasculares o de "organismos modulares".

Vakeley y Hey comparan los modelos de especiación con las secuencias de ADN. Muestran que las mutaciones deletéreas pueden verse arrastradas por otras seleccionadas. Ocurren divergencias por causa de la "contaminación" inducida por otros genes.

Amato y colaboradores comentan la replicación de las cadenas, por acción de polimerasas, dentro del complejo *Caiman crocodilus*. Vogler trata del riesgo de extinción de las especies y de su alternativa, la conservación. Acaba volviendo a la sistemática clásica, complementada por la aceptación de clinas en los haplotipos, con la posibilidad de que persistan diversos alelos en cada "variedad" (supongo que debe ser lo que ocurre con *Ostrea, Mytilus* y muchos peces).

Rootman y Cheverud abordan los *loci* que corresponden a caracteres cuantitativos y su relación con la evolución. Ochman y Grossman se refieren a la base genética de la variabilidad fenotípica en bacterias. DeSalle y Schierwater consideran la sistemática de taxones de alto nivel. Wheeler se ocupa de programación dinámica y soluciones heurísticas en la reconstrucción filogenética de secuencias moleculares, tema de interés para los taxónomos. Goldstein y Specht se centran en los riesgos de error en cuestiones, muchas de ellas de cariz filosófico, que se presentan al expresar las relaciones entre especies en forma de árboles y matrices. Larson se adentra en las dificultades específicas en la construcción de dendrogramas, procurando armonizar criterios morfológicos y moleculares. Cunningham y Collins se ocupan de la expresión de los cambios moleculares de especies marinas en un espacio de evolución. Con un ropaje nuevo reaparecen los problemas biogeográficos de siempre. El grupo formado por Jacobs, Lee, Dason, Straton y Raskoff expone la historia del desarrollo a través de la evolución molecular; llaman la atención sobre un punto sorprendente: ciertos genes relacionados con el plan general de organización, así el eje longitudinal de los animales, han persistido sin una integración necesaria en sus implicaciones evolutivas. ¿Habrá que verlo como trascendentalismo o como tipología?

Determinados aspectos de la selección natural darwinista quedan preteridos en investigaciones sobre el desarrollo. Se resiente de ello la construcción del árbol. Aunque en casi todos los procesos se puede reconocer cierta jerarquía, distintos determinantes fundamentales (carácter pluricelular, segmentación, diferenciación dorsiventral) pueden coexistir, quizás en el mismo nivel informático, con mecanismos que influyen o se expresan de manera más limitada. No debiera ello extrañarnos más que la comprobación de que un mismo lenguaje sirva para saludar al vecino y para declarar la guerra. No ha de verse como una dificultad intelectual, sino como una gracia en los caminos de la naturaleza.

RAMÓN MARGALEF

# *Seguiremos explorando los campos del conocimiento*




**LAS ESTRELLAS DE RAYOS X SUPERBLANDAS Y LAS SUPERNOVAS, por Peter Kahabka, Edward P. J. van den Heuvel y Saul A. Rappaport**

*Hace varios años se descubrió un tipo nuevo de estrella que despedía rayos X de una energía bajísima. Se cree ahora que esas fuentes "superblandas" son enanas blancas que van devorando a sus estrellas compañeras hasta que acaban por estallar.*

**EL ROMPECABEZAS DE LA HIPERTENSION EN AFROAMERICANOS, por Richard S. Cooper, Charles N. Rotimi y Ryk Ward**

*Suele culparse a los genes de la hipertensión arterial común entre los americanos de origen africano. Ocurre, sin embargo, que, en Africa, tal hipertensión no se da. Tenemos aquí un ejemplo óptimo de interacción entre genoma y entorno.*

**LA BUSQUEDA DE LA VIDA EN NUESTRO SISTEMA SOLAR, por Bruce M. Jakosky**

*Si la vida se desarrolló independientemente en nuestros planetas o satélites vecinos, ¿dónde están los lugares más probables para buscar indicaciones de organismos extraterrestres?*

**LAS AVES FOSILES DE LAS HOYAS, por José Luis Sanz y Bernardino P. Pérez-Moreno**

*En un yacimiento del Cretácico inferior de la provincia de Cuenca se está obteniendo información sobre las primeras fases de la historia evolutiva de las aves.*

**FACTORIZACIÓN DE NÚMEROS GRANDES**

**Johannes Buchmann**

*La seguridad de las codificaciones de datos se basa en la dificultad extrema que reviste la descomposición en factores primos de números con más de cien cifras. Pero las técnicas de factorización han experimentado grandes avances; por eso, los sistemas criptográficos utilizados hoy quizá dejen de ser seguros en un mañana próximo.*

**CICLIDOS DE LOS LAGOS DE RIFT, por Melanie L. J. Stiassny y Axel Meyer**

*La extraordinaria diversidad de los peces cíclidos pone en tela de juicio arraigadas ideas sobre la velocidad de los procesos de especiación.*

**LA VIA DE SALIDA AL ESPACIO, por Tim Beardsley**

*Para adentrarse todavía más en el espacio profundo, el hombre tendrá que empezar por idear un modo de acceso más económico y eficaz. Para nuestra ventura, no faltan propuestas.*

**AERORREACTORES, por Charles R. McClinton**
**CABLES ESPACIALES, por Robert L. Forward y Robert P. Hoyt**
**AUTOPISTAS DE LUZ, por Leik N. Myrabo**
**VELAS DE LUZ, por Henry M. Harris**
**COHETES NUCLEARES COMPACTOS, por James R. Powell**
**ALCANZAR LAS ESTRELLAS, por Stephanie D. Leifer**

INVESTIGACION y CIENCIA